Colour Vision Test for Railway Dispatchers

by

Shankaran Ramaswamy

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Vision Science

Waterloo, Ontario, Canada, 2009

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Chapter 1

Introduction

Colour codes are used extensively in railways to convey specific information governing movement of trains and equipment on the track. One such task is the railway traffic control display that uses colour coded video display terminals (VDTs) to convey information of the signal status, train movements and track status to the railway dispatcher. Because individuals with colour vision deficiencies (colour-defectives) may have problems with these colour-related tasks, questions were raised about the suitability of colour vision defectives to work as railway dispatchers. In order to answer that, a VDT based Dispatch Colour Vision Test based on the actual railway traffic display was developed previously.

Purpose

The main purpose of this thesis is to establish the pass/fail scores and repeatability of the VDT based Dispatch Colour Vision Test that resulted from the previous work. Secondly, the study will also examine whether clinical colour vision tests can predict the performance on the practical task.

Chapter 2, 3 and 4

Methods

The Dispatch colour vision test was divided into three parts based on the colour sets that the dispatcher had to recognize. The testing computer system used the the same RGB colour

settings, graphics card and monitor as in railway dispatch centres. Subjects viewed the display colours and entered their responses by using a mouse. One hundred colour-normals and fifty two colour-defectives participated in the initial session. The test was repeated approximately after 10 days. Ninety three colour-normals (93%) and 44 (85%) colour-defectives participated in the second session. The total number of errors and time to complete the test was recorded.

Chapter 5-9

Results

Pass/Fail on the VDT Dispatch colour vision test was based on colour-normal errors. Ignoring orange-red errors, two errors were allowed in the first session and one error was allowed in the second session. Based on this criterion, 42% of colour vision defectives could perform as well as colour normal subjects. The kappa coefficient of agreement between the sessions for the colour-defectives was 0.85.

Detailed analysis between the colour differences and the errors showed only a weak correlation between the two. However, the general trend was that colour-defectives made more errors on colours that were near or along the same lines of confusions and the colours were nearly equal in luminance. Nevertheless, the interaction between luminance and location with respect to the lines of confusion was not easy to interpret.

The time to complete the task for the colour-defectives who passed the test took 14% longer than colour-normals and colour-defectives who failed took 30% longer than colour-normals. All

groups showed a similar learning effect with an 18% reduction in mean times to complete the task at the second session. There was no significant correlation between the number of errors and time to complete or the clinical tests and completion times for any of the groups.

Clinical colour vision tests have limited value in predicting performance of colour-defectives on the Dispatch test. Logistic analysis results showed that the Farnsworth D-15 along with the Nagel was the best predictor of the VDT Dispatch colour test pass/fail results. However, these results were similar to using the Farnsworth D-15 test alone. Ninety-five percent of the individuals who failed the Farnsworth D-15 also failed the Dispatch test. However, approximately 25% of the individuals who passed the Farnsworth D-15 failed the VDT Dispatch colour test which is an unacceptable false negative rate. These results indicate the Farnsworth D-15 can only be used to predict who is likely to fail the dispatch test.

Chapter 10

Conclusions

Forty two percent of colour vision defectives could perform as well as colour-normals in identifying VDT railway display colours and time to complete the task. Clinical colour vision tests were inadequate predictors of performance in practical task, overall. However, the Farnsworth D-15 was a very good predictor of who would fail the VDT Dispatch test. Hence a practical VDT Dispatch test may be needed to test individuals who would want to work as railway dispatchers.

Acknowledgements

I would like to convey my sincere thanks to my advisor, Dr. Jeff Hovis, whose guidance, support, understanding and patience, added remarkably to my graduate experience.

I would also like to express my sincere gratitude to my committee members Dr Ralph Chou and Dr. Chris Hudson for their constant support and encouragement.

My special thanks to Dr. Thomas Singer and Dr. Natalie Hutchings for their assistance.

I would like to thank the computer technicians Jim Davidson, Chris Mathers and the graduate coordinators Sharon Dahmer, Kristin Snell and Krista Parsons for their help and support.

My special thanks to the UW OLRC library staff members for providing fantastic service and guidance when in need. They made my research life better.

I am very grateful to my fellow graduate students, staff, faculty members and landlords who provided me a very friendly and cordial environment during my stay here at Waterloo.

Finally, I would like to thank my parents, brother and relatives for their love, moral and financial support and encouragement.

Dedication

To my parents and brother for all their love, support and encouragement

Table of Contents

CHAPTER 6. Results II. Types of Errors made by Colour-Defectives in the VDT Test…. 92

CHAPTER 7. Results III. VDT Test Colour Differences and Error Correlation…..……. 120

CHAPTER 8. Results IV. Time Taken to Complete the VDT Test……….…..…………. 136

CHAPTER 9. Results V. Comparison with Clinical Tests………...………….……………… 156

List of Tables

List of Figures

CHAPTER 1

1.1. Introduction

Colour coding has played a major role in conveying information in the transportation industry for over two centuries.[1, 2] In the railway industry, coloured signals are still used to convey information to the engineer and conductor governing the movement of their train. However, the use of colour has expanded with introduction of video display terminals (VDT) into the dispatch centres. Train movement on the network is controlled by dispatchers located at different sites throughout North America. Colour-coded VDT displays provide real time information about the track status, wayside signal lights, approximate train location, and permissions for movement of maintenance equipment and personnel on track besides the train.

Although the increased usage of colours to convey information has improved efficiency and safety, it also possesses a set of challenges for colour-defective workers who may have problems in recognising certain colours correctly. About 8% of the male population and 0.4% of the female population have defective colour vision since birth. The colour vision defect remains stable throughout life.[3] The ability of these individuals to distinguish and identify colours correctly varies depending on severity and type of defect.

1.2. Congenital Colour Vision Deficiencies

Congenital colour vision defects are separated into two major classes. The division is based on whether the individuals with the defect have difficulty discriminating colours along the red-green axis of the colour circle or the blue-yellow axis. The most prevalent is the congenital red-green colour vision deficiency. This is the deficiency that corresponds to the often cited 8% of the males and 0.4% of the females. In contrast, the congenital blue-yellow defects are rare. The estimated prevalence is about 0.005% of the population.[4]

Congenital colour vision deficiencies arise from differences in the cone photopigments compared to the colour-normal population. The red-green colour vision defects occur because the photopigment in the Long wavelength sensitive cone (L-cone), or Medium wavelength sensitive cone (M-cone), or both cones are different from the rest of the population or one of these photopigments is not expressed in the retina. These defects have an X-linked recessive inheritance pattern. Individuals with a congenital red-green defect can be divided into two groups. One group is referred to as protan and the other is referred to as deutan. A protan defect occurs when L-cone photopigment is either missing or the absorption curve is shifted to shorter wavelengths relative to the colour-normal L-cone photopigment. A deutan defect occurs when the M-cone photopigment is either missing or the absorption curve has shifted to longer wavelengths. Individuals with a congenital blue-yellow defect appear to have S-cones with a nonfunctional or abnormal photopigment. Congenital blue-yellow defects are referred to as tritan and have an autosomal dominant pattern of inheritance, but with variable penetrance. That is, individuals with same genotype will exhibit variable degrees of severity.[5]

Within each colour deficient group, there can be dichromats and anomalous trichromats. Dichromats have the most severe form of the congenital defects. These individuals have only two classes of cones present in their retina, but they have same number of cones as colour-normals. A protanope appears to be missing the L-cone and behaves as if he has only M and S-cones present in his retina. Most of these individuals are missing the L-cone photopigment gene and so only the M-cone photopigment is expressed. A deuteranope appears to be missing the M-cone and behaves as if he has only L and S cones in his retina. Most of the deuteranopes are missing the gene for the M-cone photopigment and so only the L-cone photopigment is expressed.[5-7] A tritanope appears to have a non-functional S-cone and behaves as if he has only M and L cones in his retina.

Anomalous trichromats comprise the majority of congenital red-green colour defectives. They are characterized by requiring three primaries in order to make a colour match (like colour-normals), but the proportions of the primaries in the mixture are clearly outside the normal range. The majority of the anomalous trichromats also have reduced colour discrimination. As with the dichromats, there are two subcategories of anomalous red-green trichromats. Deuteranomalous individuals are characterized by requiring more of the green primary when mixed with red to match a standard yellow. For this reason, they are often referred to a “green weak”. They have three distinct cone classes, but their “M-cone" photopigment has an absorption function that is shifted to longer wavelengths relative to the M-cone found in colour-normals. This "shifted" photopigment could actually be a hybrid L-cone pigment that has an absorption spectrum that is shifted to shorter wavelength relative to the other L-cone in the retina.[5, 7, 8, 9] That is, some deuteranomalous individuals could have two slightly different L-cone photopigments in their retina.

Protanomalous individuals are characterized by requiring more of the red primary when mixed with green to match a standard yellow. For this reason, they are often referred to a "red weak". Similar to the protanopes, protanomalous individuals also have a decreased sensitivity to red lights. The "L-cone" photopigment in these individuals has absorption function shifted to shorter wavelengths relative to a colour-normal L-cone. Recent research has shown that in many cases, the anomalous photopigment is actually a hybrid M-cone photopigment that absorbs light at slightly longer wavelengths than the other M-cone in the retina.[5, 7]

1.3. Colour Specification and Lines of Colour Confusions

In order to determine whether the colours can be discriminated by a colour-defective individual, the colours are specified graphically and their locations are compared to discrimination zones called the lines of confusions. The 1931 Commission Internationale de l’ Eclairage (CIE) chromaticity diagram is often used to specify colours graphically. Figure 1.1 shows the diagram. The diagram is based on colour mixing experiments and the position of any colour within the diagram is based on the relative amounts of the three primaries used to match any colour by colour-normals. The horseshoe-shaped curve denotes all the spectral colours, starting at violet at the lower left hand corner and looping through blue, green, orange and red at the lower right hand side. The line connecting the violet wavelengths with the red wavelengths is referred to as the line of purples. The coordinates of all colours fall within these limiting boundaries.

The solid lines radiating out of the lower right corner in Figure 1.1 are the protanopic lines of confusion. Each line corresponds to the colours that require the same ratio of the two primaries that the protanope uses to match colours. This means that the colours on any single line will appear identical when the luminances are identical. Different lines represent different ratios and therefore the colours on different lines should appear different even when they are equal in luminance. In fact the distance between any two adjacent lines represents a just-noticeable difference in colour for the protanope. This area is referred to as the zone of confusion because any colours that fall within a given zone will appear identical to a person with dichromatic defect. For example, the red, green and yellow colours all fall near the same solid line so these colours will appear identical to the protanope if they are equal luminance. Figure 1.2 and 1.3 shows the deuteranopic and tritanopic lines of confusion. Again, these lines are based on their colour matches and hue discrimination. Colours lying on the same line will appear identical if the luminances are equal.

Figure 1.1 and 1.2 shows that both deuteranopes and protanopes will have difficulty discriminating between greens, yellows, oranges and reds. The number of confusion lines indicates that the protanopes and deuteranopes can distinguish only 21 and 31 distinct wavelengths, respectively. In contrast, colour-normals can distinguish 150 distinct wavelengths.[10] Figure 1.3 shows that the tritanopes will have difficulty discriminating between grey and white, grey and yellow, grey and green, green and dark-green and blue and blue-green. The number of confusion lines indicates that the tritanopes can distinguish only 44 distinct wavelengths.[11]

The orientation and spacing of the lines of confusion represents the average of several dichromats using a 2 degree centrally fixated field of moderate duration. Thus, individual performance of dichromats may vary from predictions based on the lines of confusion, especially if the field size or duration is different for the values used in Pitt’s experiment.[10]

The anomalous trichromats fall in-between the colour-normals and the dichromats in colour discrimination. Their confusion zones appear as series of ellipses (Figure 1.4) with the major axis of each ellipse along the corresponding dichromatic confusion zone, but shorter than that of dichromats as they do not include complete range of confusions as in dichromats.[12] The length of the major axis of the ellipse varies with the severity of the defect.

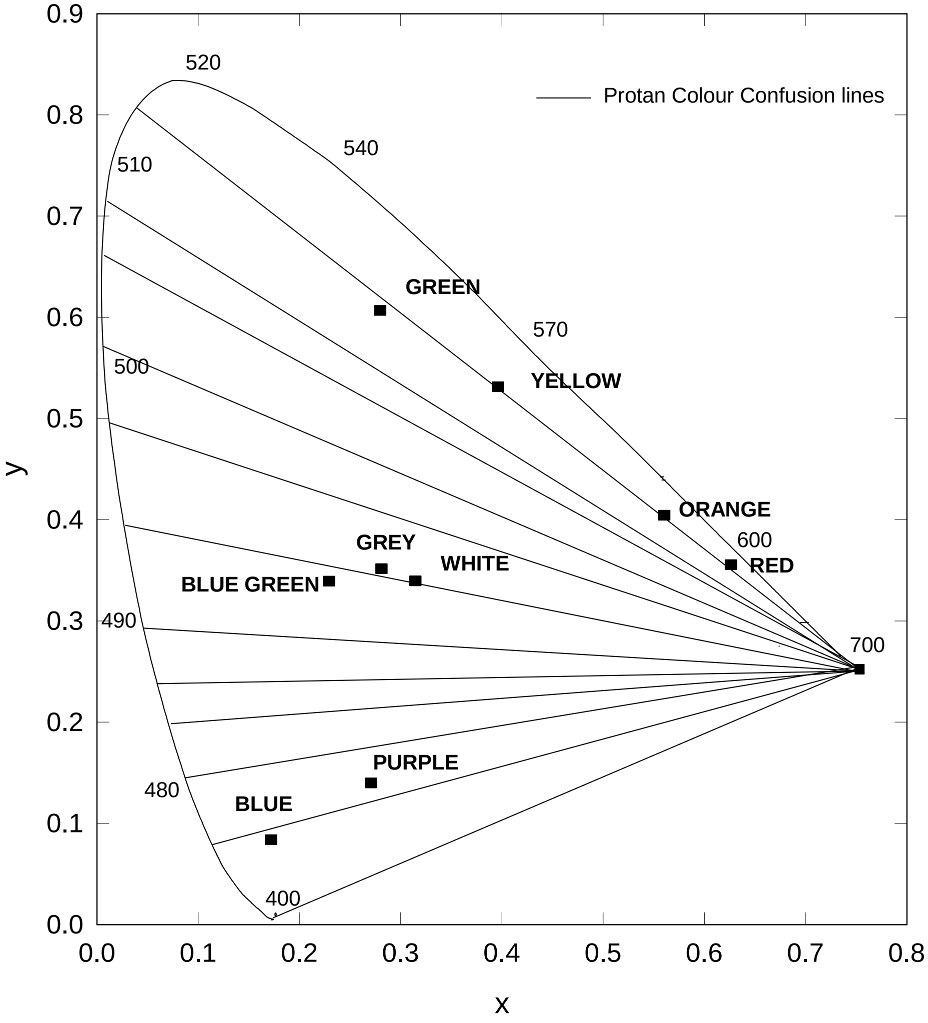


Figure 1.1. Protanopic lines of confusion in the 1931 Commission Internationale de l' Eclairage (CIE) chromaticity diagram.

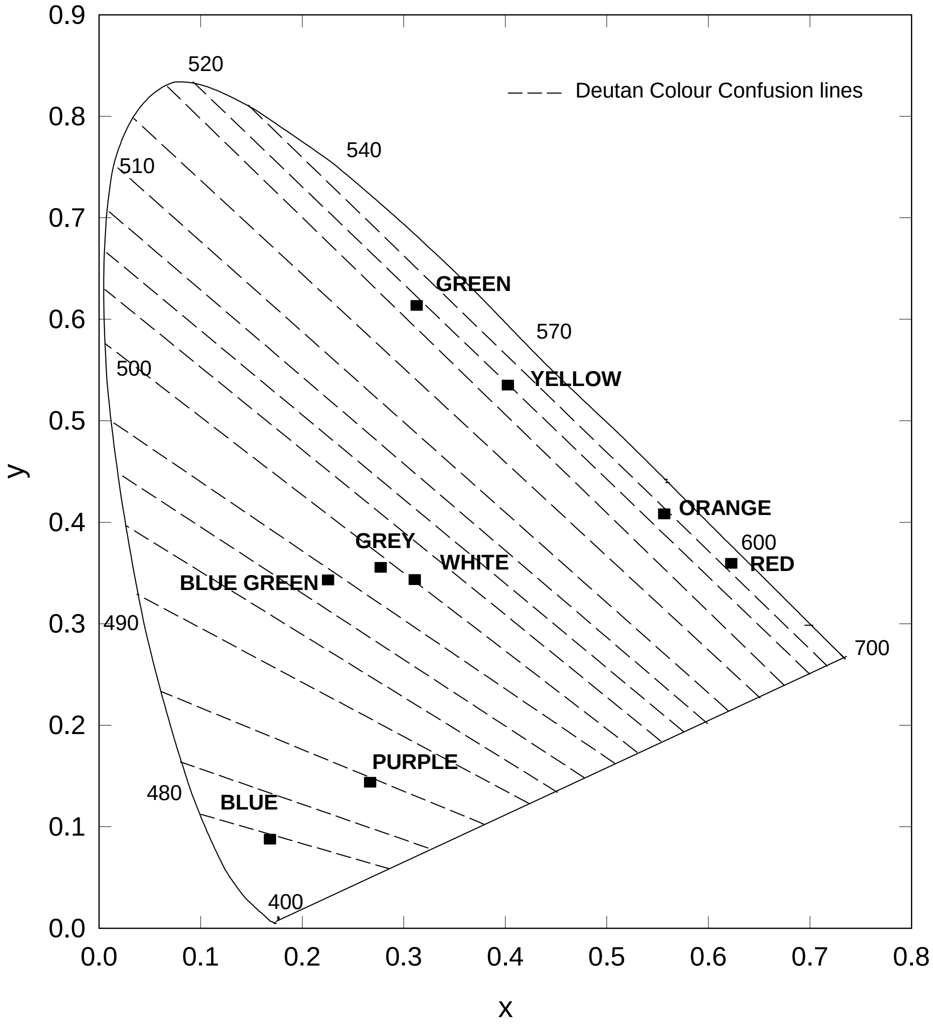


Figure 1.2. Deuteranopic lines of confusion in the 1931 Commission Internationale de l' Eclairage (CIE) chromaticity diagram.

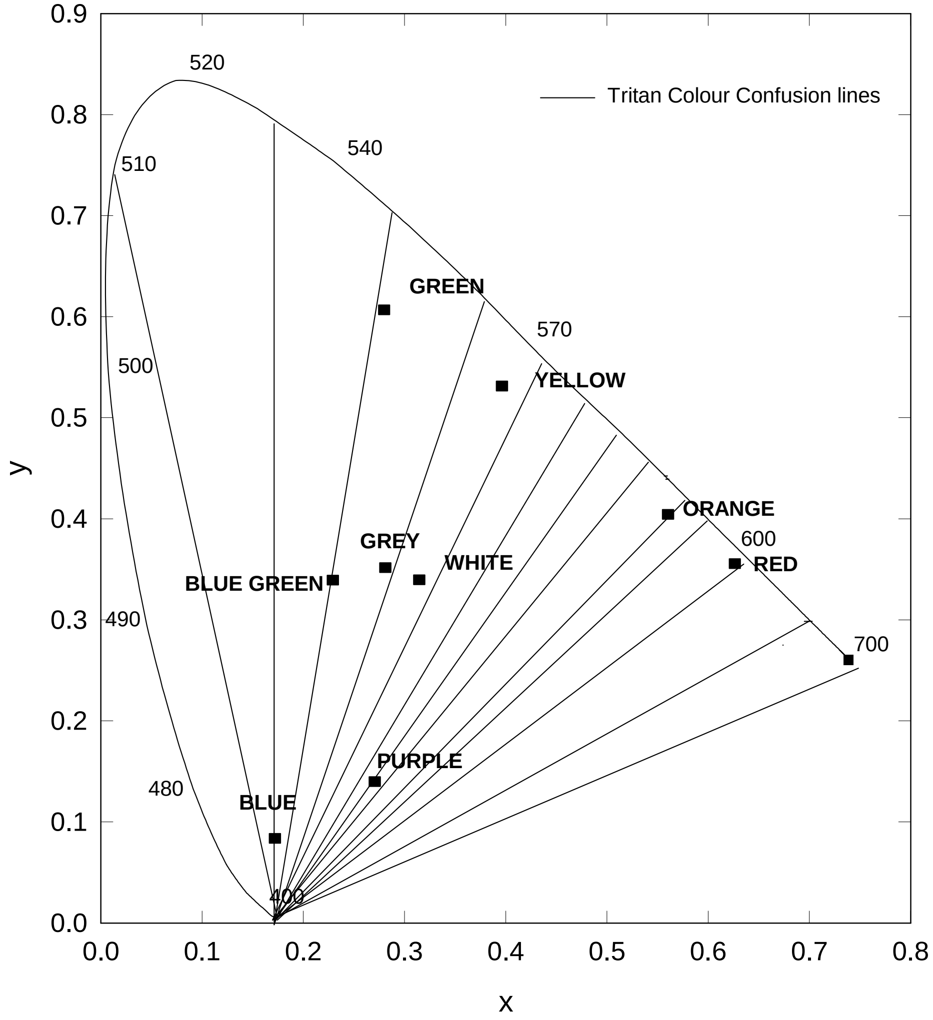


Figure 1.3. Tritanopic lines of confusion in the 1931 Commission Internationale de l' Eclairage (CIE) chromaticity diagram.

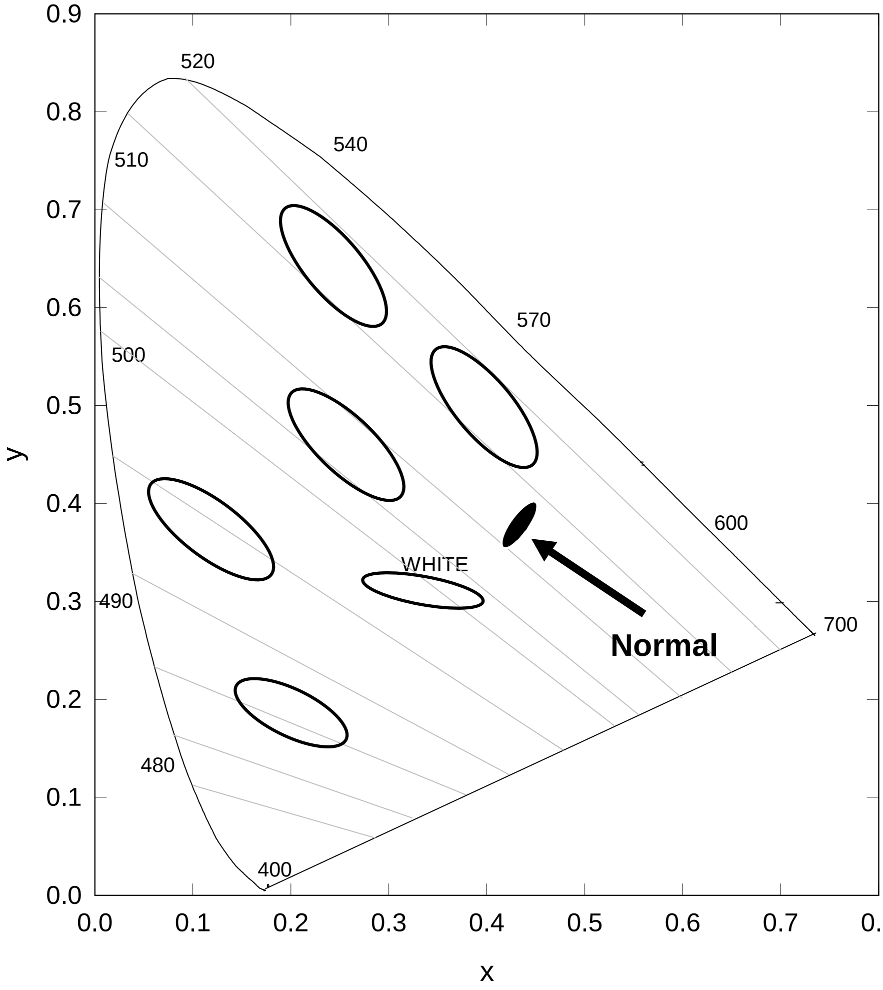


Figure 1.4. Colour confusions obtained for a deuteranomalous (Birch-Cox 1974) in the 1931 Commission Internationale de l' Eclairage (CIE) chromaticity diagram. The length of the major axis indicates the severity of the defect. The grey lines are the Deuternaopic lines of confusion.

1.4. Origin of Colour Vision Standards in Railways

Colour vision standards were the earliest form of vision standards in the rail industry. Vingrys and Cole's [2] review concluded that railway colour vision standards were suggested (in the UK) as early as 1855 by Wilson, but not adopted until 1877. Nevertheless, colour vision standards preceded visual acuity requirements by at least 25 years.[2] The need for colour vision standards arose from the realization by railway officials that there were a significant number of males with impaired colour perception and this impairment could lead to disastrous errors in the recognition of coloured signals used to control train movement. This risk was realised in 1877 when Holmgren attributed the cause for a train accident that occurred in 1875 in Lagerlunda, Sweden to defective colour vision. [13] As a result, nearly every railway company in Europe and the United States adopted colour vision standards for their employees, if they did not exist already.

The first colour vision standards appear to be based on a signal recognition performed using signal lights in the field. Testing was later done with a lantern test that simulated rail signal lights.[2] The lantern test is based on six railway signal colours like red, yellow, white, green, blue and purple. In the UK, Holmgren Wool Test was also used briefly for colour vision testing. In this, the subject is required to sort skeins of wool according to colour. As this test was not as sensitive as lantern test in identifying colour-defectives, the lantern tests were re-introduced for colour vision testing in the late 1800s and early 1900s.[2]

The ability of colour defectives to carry out tasks that involve colours has been a major concern for employers. Despite the ability of some colour defectives to carryout certain jobs,[1] the railway industry has often been stringent in their colour vision standards, essentially excluding all individuals with a colour vision defect. The justification of these rigid standards was the high cost to property or lives if an accident occurred because the locomotive engineer failed to identify the signal properly. In the dispatch center, failure to monitor train movement and track permissions could similarly result in the cost of lives or extensive property loss. However, with the emergence of equal opportunity law, the employers have an additional responsibility to justify the exclusion of colour defectives.[14] In this case, the employer must demonstrate that any given person with a colour vision defect cannot perform the job correctly and many countries require an individual assessment. This legal precedent has to lead to re-emergence of occupationally based colour vision tests.

1.5. Classification of Colour Tasks

Colour application in industry and the everyday visual environment can appear diverse and complex. However, Cole has broadly classified the applications into four categories.[14, 15] They are (1) comparative colour tasks, (2) connotative colour tasks, (3) denotative colour tasks, and (4) aesthetic colour tasks.

(1) Comparative colour tasks involve judgements of colours in terms of either matching colours or distinguishing between colours usually requiring a fair degree of precision. An example would be matching the paint on the locomotives.

(2) Connotative colour tasks use colour codes to convey specific information. An example would be the coloured signal lights used in transportation.

(3) Denotative colour tasks use colour to mark out or identify objects. An example would be using colours to facilitate visual search in complex displays.

(4) Aesthetic colour tasks use colour to create an emotional response or convey a mood. An example would be the decorative lighting used to highlight buildings.

1.6. Connotative Colour Tasks in Railways

Connotative Colour Tasks are widely used in maritime, railways, roadways and aviation to convey information.[14, 15] Traditionally, colour codes have been used in railways for long range signalling. These tasks vary from signal lights for long range viewing distances at different distances to surface colours (ie flags) with varying luminosity for short range signalling. Because of the relatively long distances (e.g. 0.5 to 1.6 km) involved in viewing the signals, conveying the visual information using colours provides more options for railways. Coloured signals are used to inform locomotive engineers about the conditions ahead and the track restrictions. Often there is little or no redundancy and the ability to recognise such colour tasks is considered critical from a safety perspective.

Other positions in railways, for example railway traffic control, have comparatively less critical colour vision demands because target size and intensity of the coloured signals is greater and there is also the possibility of redundant coding. This research deals with one such task, namely, the railway traffic control display that uses the computer based visual display units with colour codes to convey information to the railway dispatcher.

This display uses colour to code track information, signal status, train movement and rules that govern the track usage. The dispatchers in the network management centres have to be able to correctly identify the colours in the traffic control display in order to operate safely and efficiently. There is often no redundant coding in this information. Because of the diverse nature of the colour codes and the ability of some colour defectives to recognise colour codes,[14, 16] there have been concerns about the suitability of colour defectives in this occupation.

1.7. Colour Defectives and the Risk Factor

The potential safety risk is that the colours used in railway dispatch centres are also the ones that colour defective individuals are most likely to confuse. Hence, there are concerns about their ability to perform colour tasks as a railway dispatcher.

However, the presence of a colour vision defect does not automatically exclude a person from this position, since individuals with mild colour vision defects (especially deutans) can perform as well as normals on certain VDT displays that use similar colours.[16] Nevertheless, an individual's exact performance appears to be associated with the specific display characteristics such as the colour set, brightness differences between the coloured objects and the severity of the defect.

The variability of colour defective's performance becomes apparent when reviewing studies correlating performance of clinical tests with practical tests.[1, 16-19] The general finding is that mild colour defectives perform colour tasks better than severe colour defectives, but in some cases, the dichromats could perform as well as individuals with mild-to-moderate defects. There is also an interaction with the display characteristics. Colour-defectives tend to make fewer errors when the objects are bright, large in size, and the colour differences are large. Thus, many colour-defectives may perform as well as colour-normals depending on the specific display characteristics. Although there is often a correlation between the clinical test results and performance on a practical task, one cannot generally use the clinical tests to predict an individual's performance on the task because of these two factors and so a practical test which represents the actual task may be required to assess a person’s colour vision.

1.8. Establishing a Bona fide Occupational Standard.

Colour recognition problems vary with type and severity of colour deficiency along with the display characteristics. Because of this variability, employers may not necessarily exclude individuals with a colour vision defect unless they can demonstrate a risk to safety or undo hardship in terms of cost. In some cases a demonstration of risk may have to be made on an individual basis. That is, the colour-defective person may only be excluded based on his/her individual performance.[13] This means that bona fide occupationally based colour vision standards or tests often need to be developed.

In establishing a bona fide standard or tests in Canada, there are a number of steps involved [20]

1) Forming a project management team consisting of all stakeholder groups, including scientific experts, subjects matter experts (management and union), human resources, human rights and legal counsel and establishing clear objectives.
2) Job familiarization
3) Job demands review and analysis
4) Deriving a representative subset of physically demanding tasks

5) Characterisation of Representative Tasks
6) Development of Test Protocol: Job simulation tests
7) Standardization of Test Protocol
8) Establish scientific accuracy of the test protocol
9) Development of Performance standards
10) Evaluating the results of applying the standards to the performance incumbent.
11) Implementing test protocol
12) Maintain and ongoing review.

To summarise, the general recommended procedure in establishing a bona fide occupational standard is to determine the needs of the occupation, the task (colours used), the importance of colour judgements and how frequently these judgments have been made. A standard is then proposed, which can be implemented with established reliable tests. Although this means that the clinical tests may not be used to define the standard, they may still play a role in developing an efficient testing protocol. For example, a colour vision screening test is used to identify individuals who are at greatest risk in making an error in identifying colours and only those individuals who fail the screening test are further evaluated with the occupationally based tests. This process was followed in the development of the dispatch test for railways.[21]

1.9. Purpose

The purpose of this study is to establish pass/fail scores and repeatability of the VDT based Dispatch Colour Vision Test developed previously. [21] The second aim of the study is to determine the correlation between the colour differences between test colours and the test errors. The third aim of the study is to determine the time taken by the colour-defectives to complete the task and to see if there is any correlation between the test errors and time to complete the task. The fourth aim of the study is to examine the correlation between the practical test and a variety of clinical tests to determine whether the clinical tests can be used to improve the evaluation procedure. The background and details of the practical test are presented in Chapter 2. The subsequent chapters give the clinical tests, methods, test results, repeatability, comparative diagnostic test performance, performance of colour defectives and conclusions.

References

1. Voke J. Colour vision defects--occupational significance and testing requirements. J Soc Occup Med 1978;28:51-6.
2. Vingrys AJ, Cole BL. Origins of colour vision standards within the transport industry. Ophthalmic Physiol Opt 1986;6:369-75.
3. Pokorny J SV, Verriest G, Pinckers AJLG eds. Congenital and Acquired Colour Vision Defects. Newyork: Grune and Stratton 1979.
4. Delpero WT, O'Neill H, Casson E, Hovis J. Aviation-relevent epidemiology of color vision deficiency. Aviat Space Environ Med. 2005;76:127-33.
5. Wissinger B, Sharpe LT. New aspects of an old theme: the genetic basis of human color vision. Am J Hum Genet. 1998;63:1257-62.
6. Nathans J, Thomas D, Hogness DS. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. Science. 1986;232:193-202.
7. Neitz M, Neitz J. Molecular genetics of color vision and color vision defects. Arch Ophthalmol. 2000;118:691-700.
8. He JC, Shevell SK. Variation in color matching and discrimination among deuteranomalous trichromats: theoretical implications of small differences in photopigments. Vision Res. 1995;35:2579-88.
9. Shevell SK, He JC, Kainz P, Neitz J, Neitz M. Relating color discrimination to photopigment genes in deutan observers. Vision Res. 1998;38:3371-6.

10. Pitt FHG. Characteristics of dichromatic vision. Report of the Committee on the Physiology of Vision: XIV Special Report Series, London: Medical Research Council of Britain, No. 200. London: His Majesty's Stationery Office. 1935.
11. Sharpe LT, Stockman A, Jagle H, J. N. Opsin genes, cone photopigments, color vision, and color blindness. In: Gegenfurtener KR SL, editor. Color Vision, from Genes to Perception.: Cambridge University Press, Cambridge, UK, 1999:3-51.
12. Birch-Cox J. Isochromatic lines and the design of colour vision tests. Mod Probl Ophthalmol 1974;13:8-13.
13. CIE Technical Report. International recommendations for colour vision requirements for transport. 2001.
14. Cole BL. The handicap of abnormal colour vision. Clin Exp Optom 2004;87:258-75.
15. Vingrys AJ, Cole BL. Are colour vision standards justified for the transport industry? Ophthalmic Physiol Opt 1988;8:257-74.
16. Mahon L, Jacobs R. Electronic Flight Information System displays and color defective observers. Clinical and Experimental Optometry 1991;74.6 Nov/Dec:196-203.
17. Kuyk TK VJr, Lahey MA, Clark DJ. Ability of deutan color defectives to perform simulated air traffic control tasks. Am J Optom Physiol Opt. . 1987;Jan;64(1):2-10.
18. Cole BL, Orenstein JM. Does the Farnsworth D15 test predict the ability to name colours? Clin Exp Optom. 2003;86:221-9.
19. Siu AW, MK Y. The performance of color deficient individuals on airfield color tasks. Aviat Space Environ Med. 2003 2003;May;74(5):546-50.

20. Gledhill N, Bonneau J, Salmon A. Bona Fide Occupational Requirements. Proceedings of the Consesus Forum on establishing Bona Fide requirements for Physically Demanding Occupations 2000.
21. Hovis JK. Colour Vision Demands for Canadian Pacific Railways Dispatchers. Prepared for John Cutbill, Chief Medical Officer, Canadian Pacific Railway, Calgary, Alberta. 54 pgs. 2001.

CHAPTER 2

Canadian Pacific Railway Traffic Control Display

2.1. Canadian Pacific Railway Displays

The dispatchers in the Canadian Pacific Railway network use two different Video Display Terminal (VDT) displays to monitor and control train movement. One is referred to as the Occupancy Control System (OCS) and the other is referred to as the Centralised Traffic Control (CTC) [1]. The OCS displays the different authorities issued for a given section of track but does not indicate the exact location of the work crew or train. Authorities are sets of rules that govern the movement of work crews or a train in a particular section of the track. The CTC provides the real time location of the train, wayside signal codes and authorities. Both systems use extensive colour codes in the connotative mode. Colour codes are also used in text display and radio communication monitors in the network management centre, but in denotative mode.

Although colour coding is used in the OCS, colour is redundant to the displayed text information and error checking programs so that adequate colour vision is not a prerequisite for this position. On the other hand, the colour coding used in CTC display is non redundant and there are fewer error checking routines programmed into the system. These two characteristics of the system, along with the fact that the colours used in the CTC display are ones typically confused by individuals with congenital red-green defects, mean that the operator must have adequate colour vision in order to carry out the work safely and efficiently. Because interpreting colour-coded information from the

CTC is the most critical colour-related task for the dispatchers, the practical test was designed based on this display.

2.2. Centralised Traffic Control Display

An example of a CTC display is shown in Figure 2.1. The small triangular icons indicate the status of the wayside signals. The colour codes are summarized in Table 2.1. The signal icons could be either flashing or on continuously. When the signals are flashing, the icon is actually alternating between one of the colours and grey.

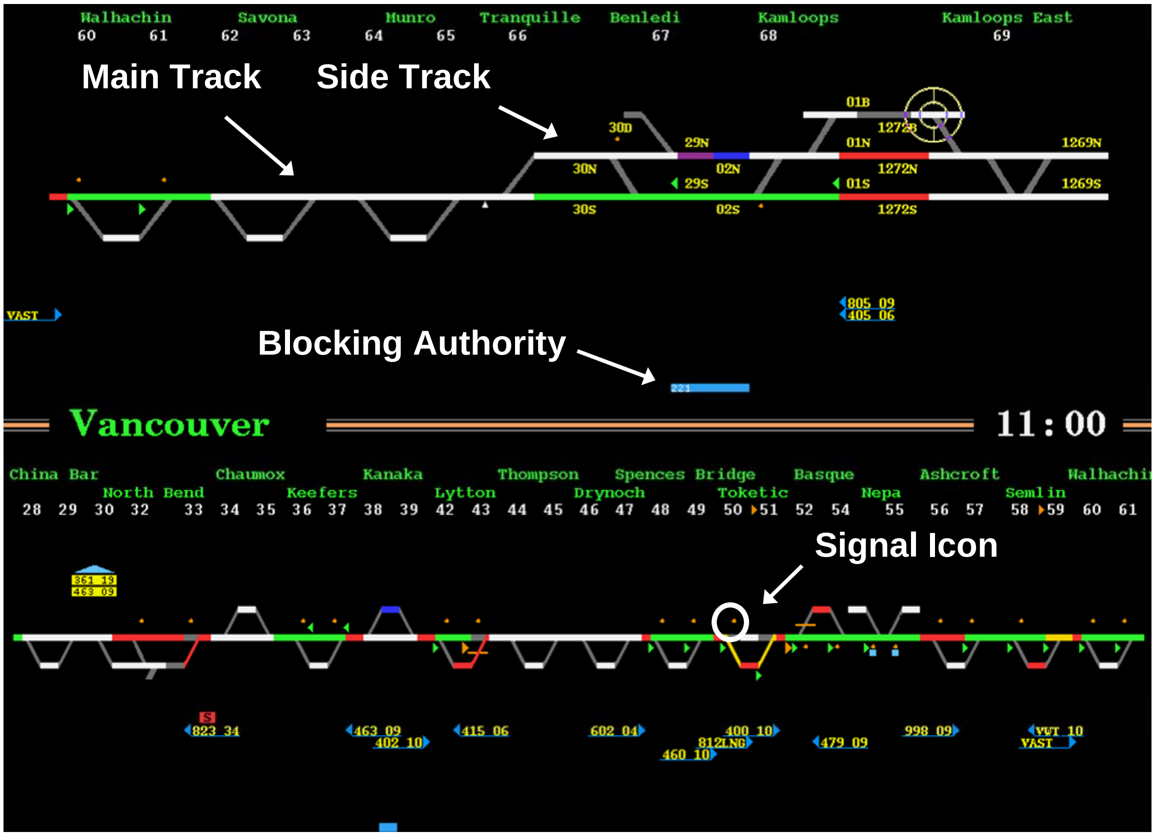


Figure 2.1. Example of a CTC display. There are two separate sections of track under control dispatcher's control. One above the "Vancouver" label and one below the label. The names displayed in green are stations along the track. Yellow numbers with the blue arrows are different trains and their direction of travel. The black font text in the yellow rectangles is a switch label. Authority codes, track graphics and signal icons are labelled.

Flashing indicates that a request has been sent to change the signal light colour, but the change has not been confirmed by the feedback system. After the change has been confirmed, the icons are on continuously and the colour is either red or green. Distinguishing between flashing red and flashing green and flashing orange is critical; however, distinguishing flashing orange from either flashing green or flashing red is aided by the fact that a horizontal orange bar, which indicates the direction of the track switch, always appears with the orange flashing icon.

Table 2.1. Colour code for signal icons (triangles).

Colour	**Meaning**
Green (Solid only – not flashing)	Confirmed Clear-Proceed
Red (Solid only – not flashing)	Confirmed Stopped
Black (Invisible to dispatcher)	Confirmed Stopped and/or Stopped and Signal turned off.
Blue	Block Confirmed -appears only on cold start of system. Changes to Red once the system is completely operational.
Flashing Red (alternating with grey)	Stop Requested but not confirmed
Flashing Green (alternating with grey)	Signal Clear requested, but not confirmed
Flashing Dark Blue (alternating with grey)	Block or Unblock Requested but not yet confirmed.
Flashing Orange (alternating with grey)	Stacked Signal-more than one signal light on display.
Grey	Alternate for flash state

Table 2.2 summarises the track status colours. Except for the yellow, there is no redundancy associated with any of the colours so that adequate colour vision is critical. Yellow indicates that the section of track is awaiting confirmation that the signal and/or switch have been aligned properly. When the section of the track is yellow, the signal icons for that section are flashing.

Table 2.2. Colour codes for track graphics (includes mainline and sidings and yards)

Colour	**Meaning**
Red	Track Occupied or Damaged
Yellow	Signal Clear Request or Stop Request Pending
Green	Clear-proceed
Blue-Green	Tracked blocked with equipment cleared to move in this section
Blue	General Track Blocking: Blocking/Unblock setting is pending as the track is undergoing maintenance or construction
Purple	Track Occupied or damaged with blocking issued for same block(s), but there is movement of equipment within that section of track
White	No activity, but that the switch is set to route traffic on that section.
Grey	Yard track & Off Route Position of Switch

Movement into and through sections is governed by the various rules called authorities. These are also colour-coded and are displayed below the corresponding section of track.

Table 2.3 summarises the authority colour codes. There are subroutines to check for errors to ensure that an authority which is entered does not contravene one that is already established. The dispatcher is also required to record the authority, when it was issued and when it was removed in a daily journal.

Table 2.3. Colour codes used for blocking authorities

Colour	**Meaning**
Red	Exclusive authority which permits movement in either direction
Yellow	Stop and proceed after permission from RTC
Orange	Joint work authority which permits more than one train movement in a specific limits.
Green	Signal/permission to enter main track.
Dark Green	General Bulletin Order (GBO) Block
Light Blue	Track occupancy permit (TOP)
Blue	Manual track block
Grey	Train is permitted in the block as well as for the maintenance of the track section by the work crew.

The dispatchers need to know where the trains are located at all times and must be able to identify the arrow icons correctly; therefore, confusion of the red with green would be a serious problem as the dispatcher may not be able to confirm the colour of the arrow icon which indicates the state or pending state of the signal. The dispatcher must also be able to distinguish red from green in order to monitor the position of the train or detect when the track is broken. They must be able to reliably distinguish purple from either blue-green or blue because these three colours code for equipment moving within a blocked section of track. Although, identifying yellow on the track grid is not as critical because the arrow icons for that section of track are flashing, distinguishing between yellow, green and red is important because the communication status of the stations is colour coded. Green lettering of the station name indicates that the settings are normal, yellow indicates that status is being checked and red coding for a failure or that control is local or given to a technician.

Initially the difficulties in identifying the authority colour codes appeared to be a less of an issue because of the error checking programs, availability of text information on other monitors and the dispatcher would hopefully remember which authorities he or she issued. However, this means there can be no room for memory lapse or distraction and the communication between dispatchers at shift changes has to be accurate. Given the number of authorities issued per day, subject matter experts state it is practically impossible to scroll through the whole text information every time an authority is issued. Although the error checking programs would most likely catch mistakes, trains could be delayed as the dispatcher tries to resolve the conflict in authority codes. Hence, despite the apparent redundancy, dispatchers still require adequate colour vision to identify the authority colours for safe and efficient operation of equipment on the tracks.

2.3. Colour Confusions of the Dispatch-VDT colours

As the blue-yellow defect is extremely rare and since it's highly unlikely that these individuals would apply for a position as a railway dispatcher this study will focus mainly on red-green defects. Figures 2.2 and 2.3 summarize the potential colour confusions in the 1931 Commission Internationale de l' Eclairage (CIE) chromaticity diagram for the display colours measured at the worksite. [1]

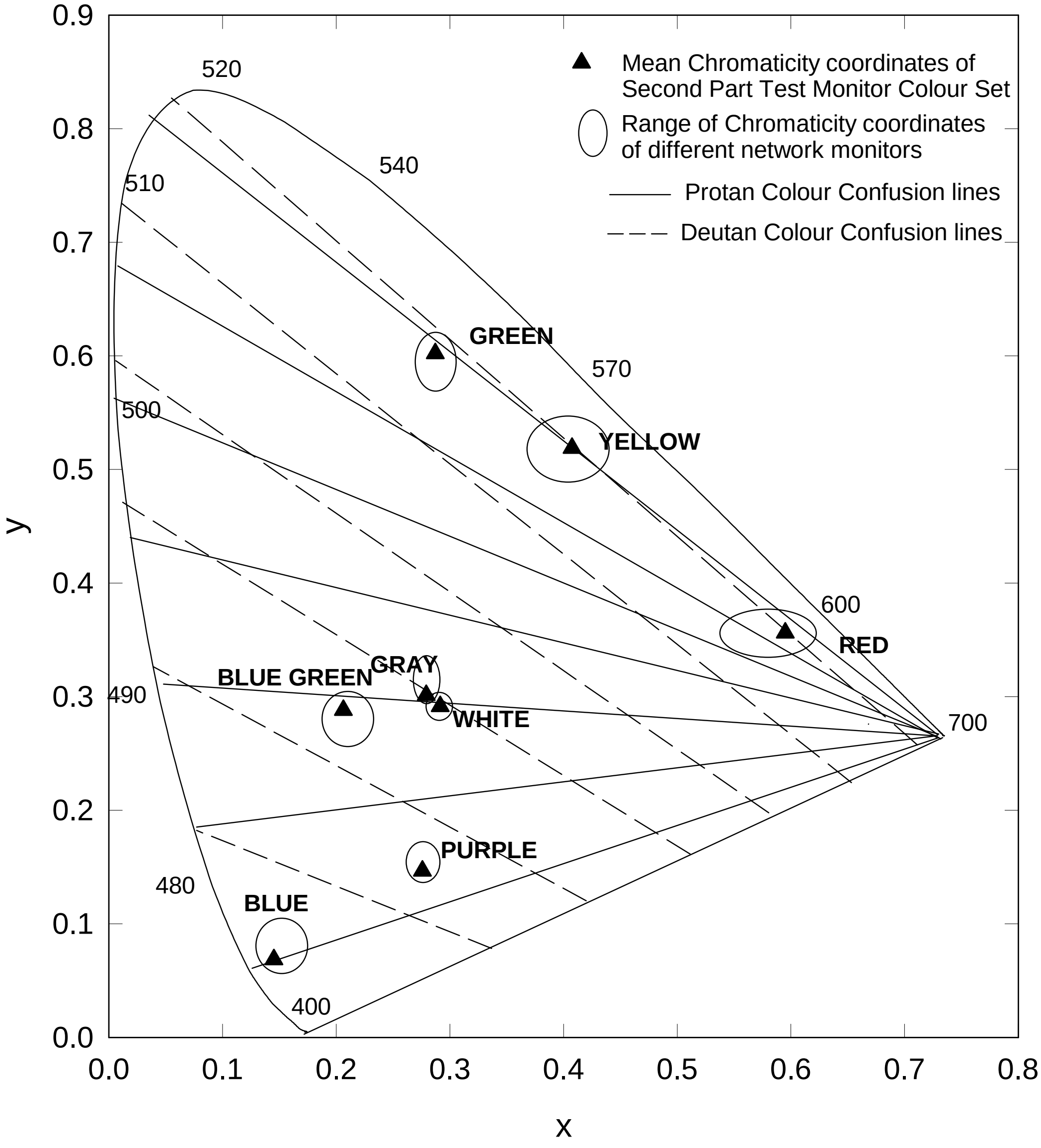


Figure 2.2. Chromaticity coordinates for the signal and track status colours in the 1931 CIE chromaticity diagram. The ellipses represent the range of chromaticity coordinates measured for various monitors at the Network Management Centre. Filled triangles are the colours used in the actual practical test. The solid and dashed radiating lines are representative protanopic and deuteranopic lines of confusions.

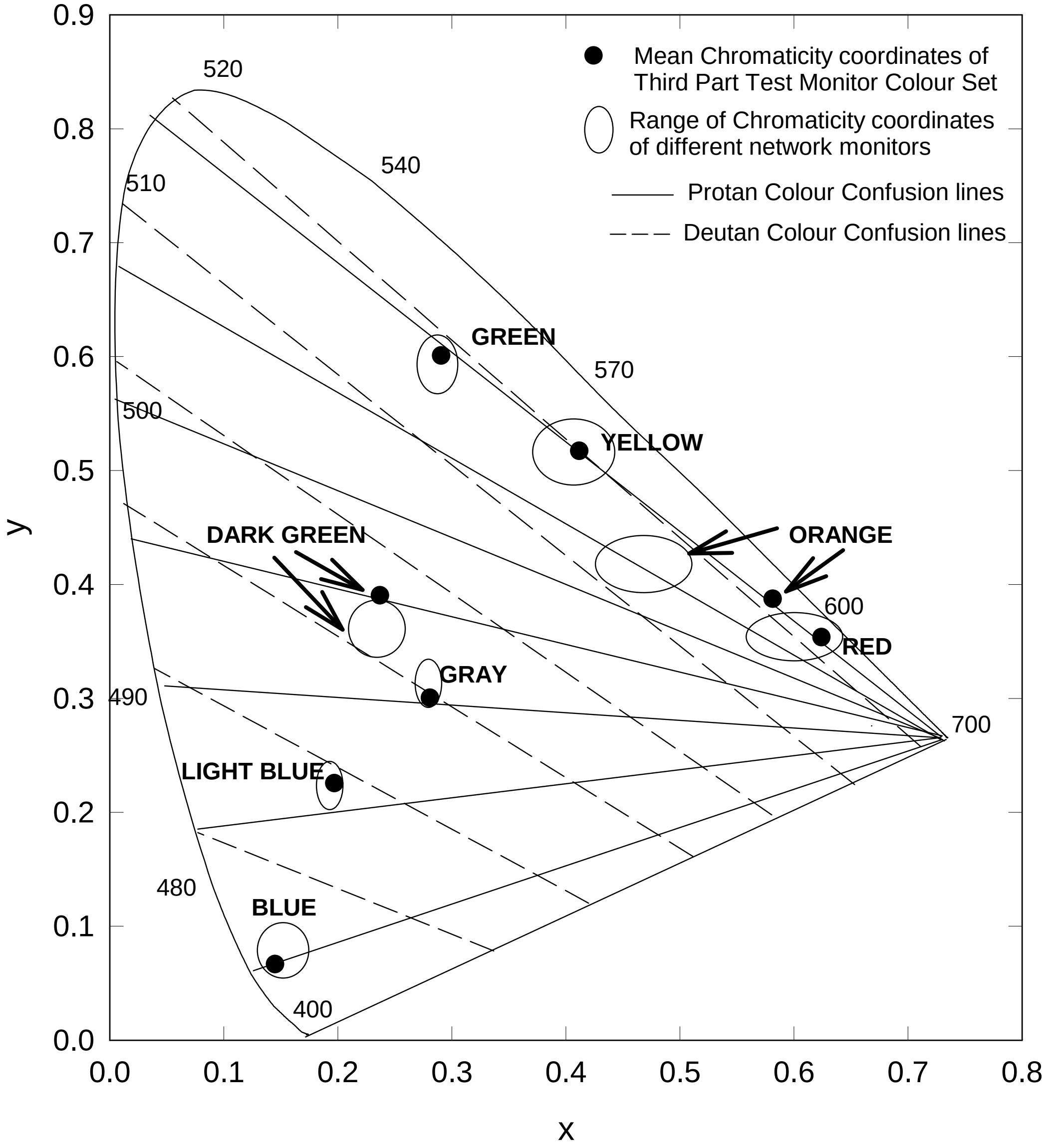


Figure 2.3.Chromaticity coordinates for the blocking authority colours in the 1931 CIE chromaticity diagram. The ellipses represent the range of chromaticity coordinates measured for various monitors at the Network Management Centre. Filled circles are the colours used in the actual practical test. The solid and dashed radiating lines are representative of protanopic and deuteranopic lines of confusion.

The figures show that the protanope is likely to confuse, red, orange, yellow and green colours with each other since they fall very close to the same line of confusion. Other sets of colours that the protanope may confuse are purple with blue, dark green with grey or white, and perhaps light blue with either white or grey. Deuteranopes may also confuse red, orange, yellow, and green with each other. Other colours that they may be confused are dark green with grey or white and purple may be confused with blue green.

The potential confusions by the deuteranopes and protanopes have certain qualifications. First, the lines of confusion in Figures 2.2 and 2.3 assume that the colours on the monitor are equal in brightness; however, Figure 2.4 shows that the luminances are not equal.

In particular, red is always dimmer than green, yellow and orange. This brightness difference between red and the other colours is even larger for the protanope who has a decreased sensitivity to long wavelength light. Similarly blue-green is always brighter than purple which may help the deuteranope distinguish between these two colours. Although there are small differences in brightness for the dark green and grey and blue and purple, it’s not clear as to whether the differences are large enough to reduce the number of confusions between dark green and grey or between blue and purple.

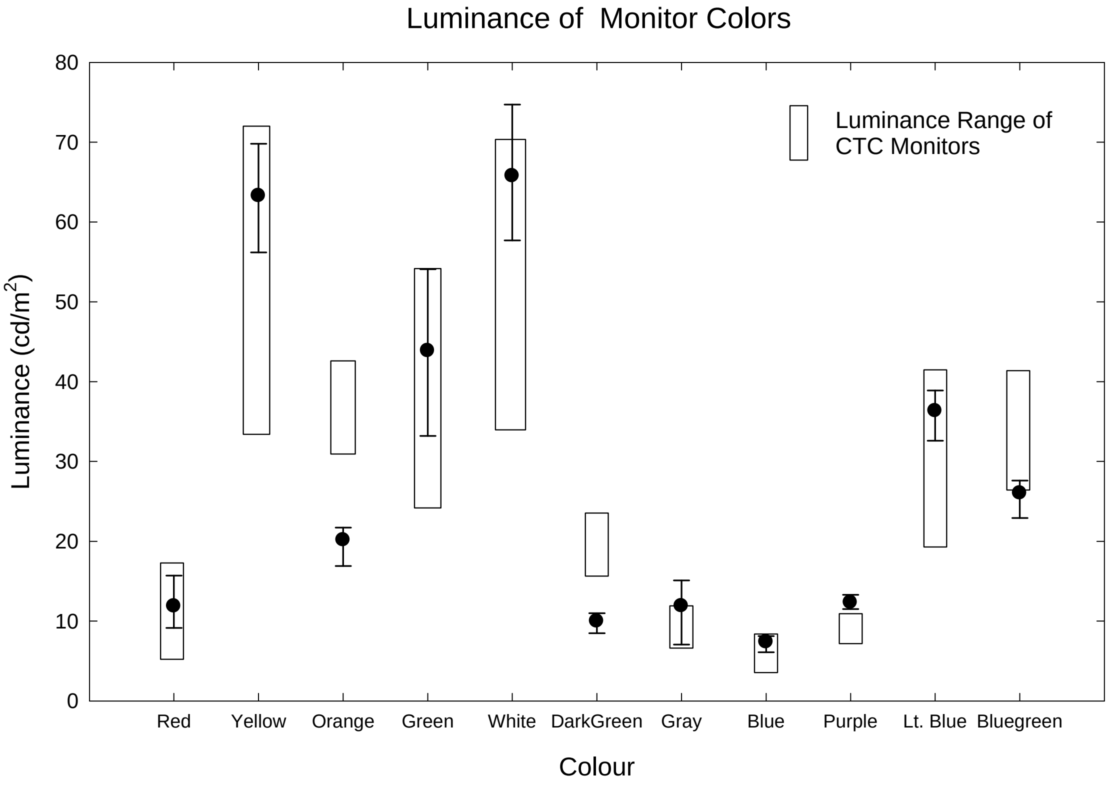


Figure 2.4. Mean luminance of the test monitor colour display colours (solid filled circle), upper and lower range of test colour luminances (error bars) and the range of luminances (rectangles) measured for various monitors at the Network Management Centre.

The second qualification is the size of the target. The smallest target measured on the display was 4 mm diagonally and the largest measured was 70 mm in length with a width varying from 2 to 3mm.[1] At a typical viewing distance of one metre, this results in objects that have an angular subtense between 14 min of arc and 2.0 degrees. The lines of confusion were obtained for 2 degree field, but for objects smaller than this (example signal icon whose side subtends 14 min of arc), one might expect the colours to be more difficult to identify for colour defectives. Cole's experiment [2] showed that surface colour objects that subtend angle less than 0.5 degrees have a more pronounced increase in errors for the colour defectives. For colour objects larger than 0.5 degree the errors were lower than smaller stimuli targets and independent of target size.[2] These results suggest the signal icon targets would be difficult to identify for colour defectives.

The third qualification is that the lines of confusion in figure 2.2 and 2.3 represent an average protanope and average deuteranope. The orientation of the lines of confusion may vary across individuals depending on the ocular media transmission and photopigments present in the retina. Furthermore, the majority of anomalous trichromats may not confuse some of the colours as their discrimination ellipses are smaller than dichromatic confusion zones.

Given the individual variations in colour vision defects, size of the stimuli and brightness differences, a practical test using the same colours as the display should establish a persons ability to identify these colours accurately. The next section describes the practical test for the dispatchers.

2.4. The VDT Based Dispatch Test

The VDT Dispatch test (VDT test) is a practical test based on the CTC display. It was designed to use the same computer colour settings, graphics card, RGB settings and monitor currently in use at the dispatch centre. The practical test was generated by a Visual Basic program written by CP personnel.

The test is divided into three parts based on the different colour sets that dispatchers must identify. The first part evaluates a person's ability to identify twenty equilateral triangles as either red or green. These triangle icons are used to code the wayside signals displayed adjacent to the track grid on the CTC display. The angular subtence of the triangle's side is 14 min of arc. This part of the test consists of two screen images, each screen displaying ten triangle icons. The colour of the triangle is determined using a random block design. Figure 2.5 shows an example of one of the screens. The subject is required to identify the colours by clicking on the circle or colour name next to the appropriate name.

Figure 2.5. Example of the signal icon section of the practical test.

The second part evaluates the subject’s ability to identify colours used in the track graphic display. There are eight coloured rectangle icons; red, yellow, green, blue-green, dark blue, purple, white and gray. These colours code the different activity of the track (Table 2). The angular size of the rectangle is 35 X 11 min of arc. This part of the test displays three different screens in sequence. Each screen has sixteen rectangles arranged in a 2 X 8 matrix. The colour of the rectangle on each screen is determined by a random block design. Figure 2.6 shows an example of one of the screens. The subject is required to identify the colour of each rectangle by clicking on the circle or colour name next to the appropriate name beneath the coloured rectangle.

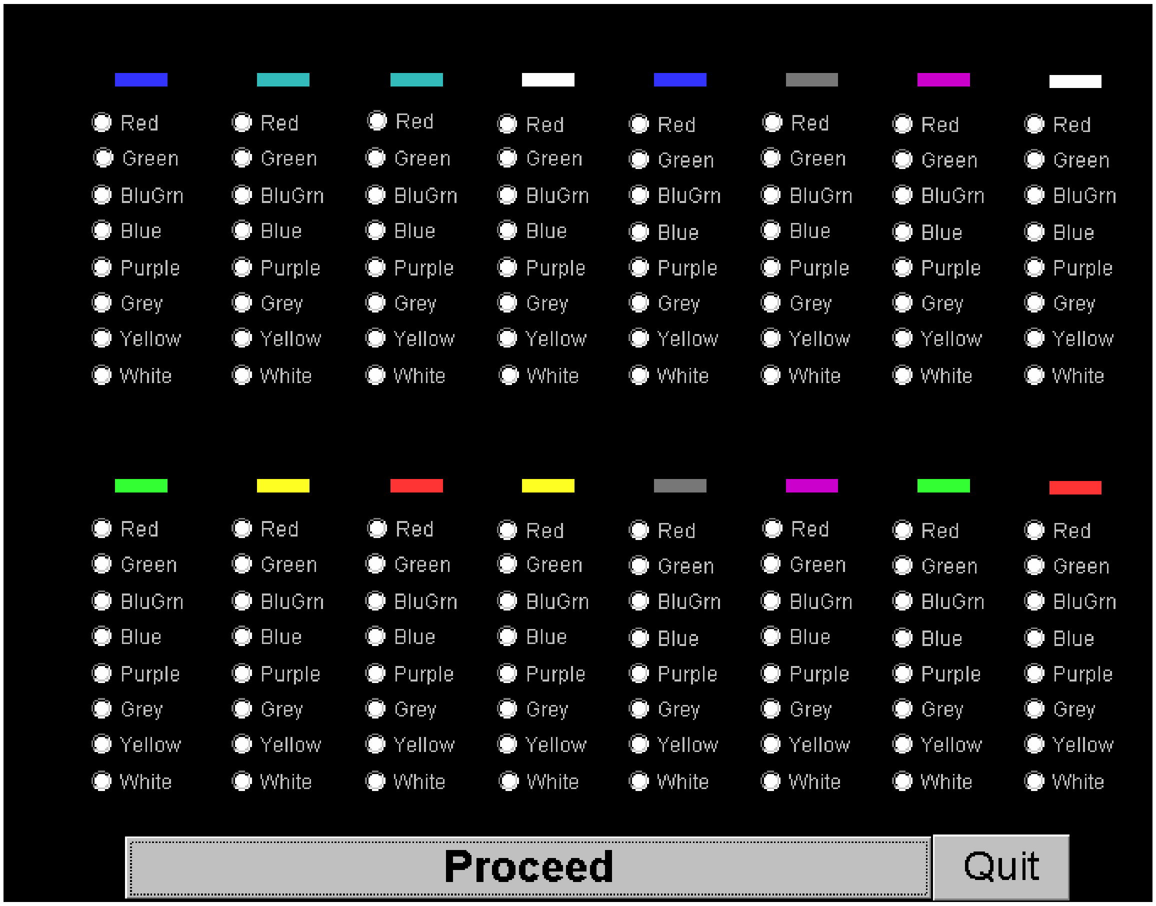


Figure 2.6. Example of the track graphics section of the practical test.

The third part evaluates one’s ability to identify the colours used for the blocking authorities (Table 3). Because the dimensions of the authority rectangles and track sections are similar in the actual display, the general design of the third part of the practical test is the same as the second part, except that the colour set is slightly different. The colours in this section are red, yellow, orange, green, dark green, light blue, dark blue and grey. Figure 2.7 is an example of one of the screens. Despite an overlap of colours between the second and third parts of the test, the two colours sets are not combined because some colour confusions that occur in the combined set are not possible in practice.

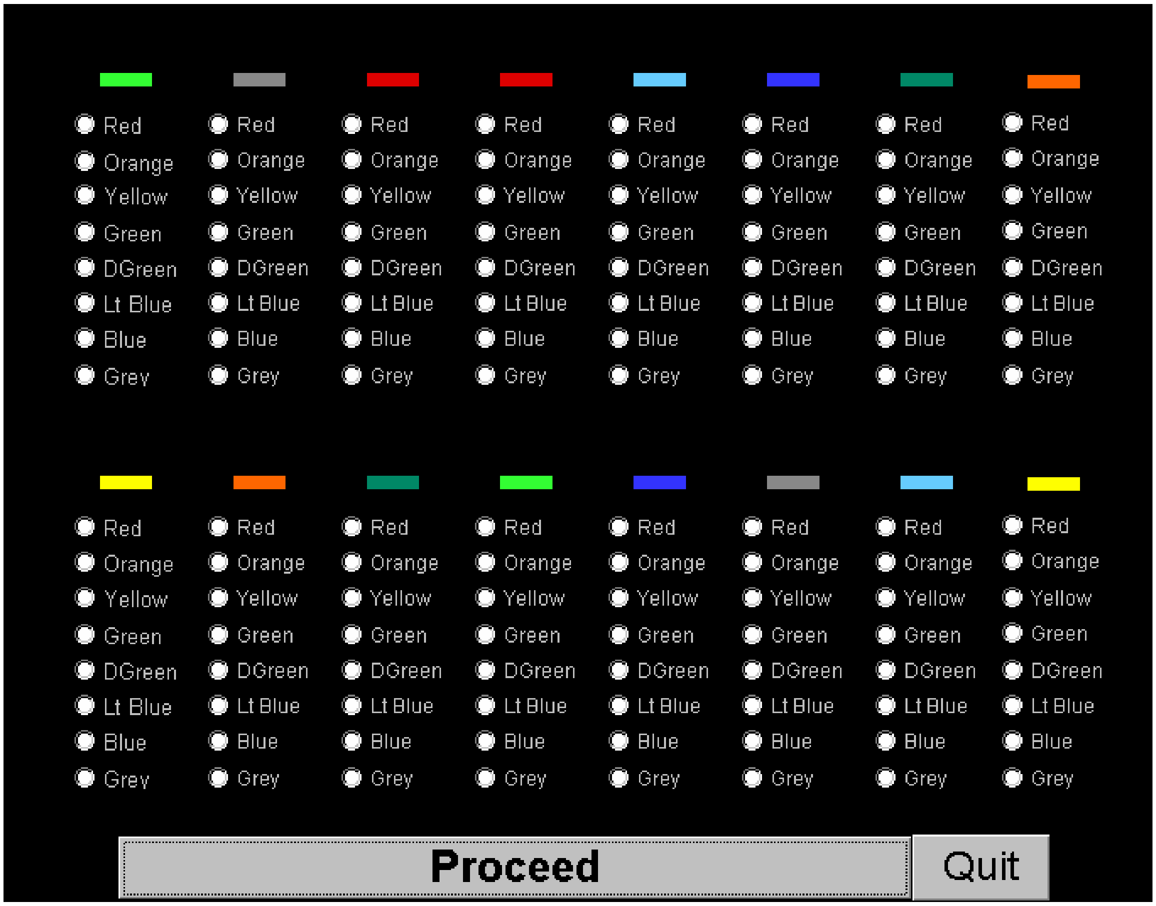


Figure 2.7. Example of the blocking authorities section of the practical test.

2.5. Test Parameters

The monitor used for this study was a 21 inch Trinitron^{TM} manufactured by Dell^{TM}. The specific settings for the monitor and graphics card are outlined in Appendix 2.1. The computer was a Toshiba PC^{TM} with a Pentium^{TM} processor and Windows 95^{TM} operating system. All the coloured icons were displayed within a dark background. The colours were measured with Minolta CS-100 Chroma meter (with a 10x close-up lens).

Figure 2.4 shows the luminances for the various colours used in the CTC display monitors. The filled solid circles are the averages and the upper and lower error bars indicate the range of luminances of the test monitor colours. The rectangle indicates the range of luminances measured at the original CTC monitors at the network management centre.

Most of the Dispatch test colour luminances fell within the range of luminances measured at the network center. The exceptions were orange, dark green and purple. The orange and dark green was dimmer whereas the purple was brighter than the on-site values. The difference in luminance between the dispatch purple and the range of actual values was small and could be considered as negligible. The difference between the luminances for the orange and dark green was more of a concern; especially since the computer system and monitor used for the Dispatch test was identical to the ones used in the CTC. One reason for the discrepancy could be the contrast and brightness adjustments individual dispatchers made to their monitor combined with the fact that the orange and dark green was displayed infrequently at the center when the measurements were taken.

The luminance range for the orange and dark green for the actual displays represent data from only one monitor at the center. Nevertheless, the relationship between the luminance of the orange and the red and green test colours was similar to trend of the actual displays. The luminances of the test dark green and grey colours were similar which was a trend found within the individual displays at the center even though the luminances varied across different displays. Although there may not be an exact luminance match between all the Dispatch colours and the actual displays, the brightness relationships between the colours were similar. This is a more important parameter to meet since subjects are likely to make judgments based on relative brightness differences rather than absolute values of the individual colours.

Figure 2.2 and Figure 2.3 show the chromaticity coordinates of the Dispatch Parts 2 and 3 test colours along with the range of values measured at the Railway Dispatch Centre.[1] Since the chromaticity coordinates of the Part 1 test colours (red, green) fell within the range of the red and green in Parts 2 and 3, they are not plotted separately. The figures show that chromaticity coordinates of almost all test colours fall within the range of values measured across different displays in the network centres. The two exceptions are orange and dark-green. The orange test colour was closer to the red than measured on the actual displays. The dark-green in the test was just near the measured range. These two discrepancies were probably due to the small number of monitors measured at the dispatch center along with slight variations in the graphics card and variations in the brightness and contrast settings of the monitors.

Generally, the chromaticity coordinates and luminances of the Dispatch test were within the range of values measured at the actual work site. This demonstrates that the Dispatch test has face validity with the actual task. The one exception is the orange test colour. Because the Dispatch orange is more similar to red, one can anticipate that there would be more errors on the orange and red colours which may have to be considered when developing the pass-fail score.

The next chapter summarises the different clinical colour vision tests used in this study.

References

1. Hovis JK. Colour Vision Demands for Canadian Pacific Railways Dispatchers. Prepared for John Cutbill, Chief Medical Officer, Canadian Pacific Railway, Calgary, Alberta. 54 pgs. 2001.
2. Cole BL, Lian KY, Lakkis C. Can color vision defective subjects who pass the Farnsworth lantern test recognize surface color codes? Aviat Space Environ Med. 2007;78:21-5.

CHAPTER 3

3.1. Clinical Colour Vision Tests

A variety of clinical colour vision tests were selected to be compared to the VDT test. This chapter would give a brief outline of those tests, their background, their procedure and the reason for their selection.

Clinical tests are often inadequate predictors of certain practical colour tasks. Despite this general finding, we wanted to know how well the clinical colour-vision tests can predict performance on the VDT practical test. Although the correlation may be less than perfect, the clinical tests may offer some efficiency in the evaluation process. For example, if all individuals who pass a certain clinical test always pass a practical evaluation, then the clinical test could be used to reduce the number of applicants who need to take the practical test. The other reason for looking at the relationship between the clinical tests and practical test is that results may provide clinicians additional information for testing and counselling colour deficient patients who use computer displays in their daily activities as to whether they will have difficulties in distinguishing between certain colours on the monitor.

The clinical tests selected for comparison were Ishihara test (38 plate edition), Nagel anomaloscope, HRR pseudisochromatic plates (3rd Edition), Farnsworth D-15, Adams Desaturated D-15 and the CN Lantern test.

3.1.1. Ishihara Test

The Ishihara pseudoisochromatic plate test is the most commonly used screening test for congenital red-green colour deficiency. The 38 plate edition has been shown to be one of the most efficient tests available to screen for red-green colour deficiency.[1-6] The Ishihara only tests for red-green defects and not tritan defects. The 38 plate edition has twenty five plates that have numerical test figures and 13 plates which test colour vision using traceable paths.

There are five different design formats within the test: [2]

1) Introductory or demonstration -Plate 1 is (numerical) and plate 38 (pathway). The figure should be identified correctly by all observers who have a visual sufficient to resolve the figures.
2) Transformation- Plates 2-9 (numerical), plates 34-37 (pathways): One number is seen by colour normals and a different or no number is seen by colour defectives.
3) Vanishing-Plates 10-17 (numerical), plates 30-33 (pathways): A number is seen by colour normals and no number is seen by colour defectives.
4) Hidden digits-Plates 18-21 (numerical), plates 28-29 (pathways): A number is seen by colour defectives and no number is seen by colour normals.
5) Classification-Plates 22-25 (numerical), plates 26-27 (pathways): The number on one side of the page is designed to be invisible to protans, but visible to deutans whereas, the other number on the page is designed to be invisible to deutans and visible to protans. The saturation of the numbers varies on different

pages. These plates allow one to classify the defect as either protan or deutan and grade the severity of the defect.

Various pass/fail criteria can be used for the test. According to the accompanying instructions, if 17 or more screening plates(less than 4 errors) are read correctly, the colour vision is regarded as normal. If 13 or less have normal responses (more than 8 errors) then the colour vision is defective. In case the subjects reads between 14 to 16 plates correctly (between 5-7 errors), then the diagnosis is considered uncertain and the patient should to be assessed with other colours vision tests including anomaloscope.

3.1.2. Nagel Anomaloscope

The Nagel anomaloscope was introduced in early 1900's and is the standard reference test for identifying and diagnosing red-green colour deficiency. The Nagel anomaloscope presents two halves of a 3 degree circular bipartite field which is viewed through a telescopic system. The bottom half of the field is a monochromatic yellow (589nm) light and the top half is a mixture of monochromatic red (670 nm) and green (546 nm) wavelengths.[4] There are two knobs on each side. One controls the brightness of the yellow and the other, the mixture of red and green mixture. The knob settings are displayed on separate scales which range from 0 to 73. The normal settings when the two halves match in colour and brightness are near 40 units on red-green mixture scale and 15 units on the yellow brightness scale. Colour-normals make a colour match within a small range of red-green mixture ratios. However, most colour defectives have match setting and range outside the colour-normal values.

Procedure: The test is done in darkness or semidarkness. The examination starts with a 30 second preadapation to the lighted Trendelburg screen located below the eye piece. After the preadaptation, the subject is asked to view the matching field monocularly with their preferred eye. The examiner pre sets a red-green mixture between 35 and 45 units on red-green scale and the yellow at 15. He asks the subject to comment on whether the two fields look identical in hue. The normal subject and the dichromat will report that the colours look identical or very similar. The anomalous trichromat will usually say that the mixture field appears too red or green. If the colours are not identical, then the subject is asked to adjust the red-green knob to make a hue match.

After making a hue match, the subject is asked if the top and bottom halves of the field are the same brightness. If not, they are asked to adjust the brightness of the yellow standard on the bottom until there is a brightness match. Next, the subject refines the hue match and the brightness match until the two fields look identical to them. This setting is the match setting and the procedure is repeated at least 3 times.

The next step is to find the range of acceptable red-green matches. The experimenter adjusts the red-green mixture +/-5 units from the median value. The subject refines the brightness match by varying the intensity of the yellow standard as needed. If the two fields appear identical, then the experimenter brackets in progressively larger steps until the extent of the range are found. If the two fields do not appear, identical, then the experimenter brackets in progressively smaller steps until the extent is found. The extent of the matching range shows the severity of the defect.[7] The subject is instructed to view the white adaptation light every 10 to 20 sec during these settings to avoid chromatic adaptation affects. Some colour defectives show an increase in their range of acceptable matches if they stare at the stimulus field for a period of time.[7]

Protanamalous individuals make matches that are usually more red than the colour normals, whereas the deuteranomalous matches usually have more green than normals. The exception would be individuals with more severe defects who have a large range of acceptable matches which include the normal settings. . Dichromats (protanopes and deuteranopes) are the extreme examples of the latter. They will accept all red-green settings as match to yellow if the brightness's are identical. This is because they have no hue discrimination for wavelengths longer than 540nm. The protanope can be distinguished from the deuteranope by the yellow brightness setting that matches the 670 nm light. Colour-normals and deuteranopes will set the yellow near 20 units, whereas protanopes will set the yellow brightness at values less than 10. This low brightness value corresponds to the loss of sensitivity to red wavelengths.

3.1.3. HRR pseudoisochromatic plates

The HRR plate test is a pseudoisochromatic design screening test designed to identify protan, deutan, tritan (blue-yellow defects based at the post receptor level) and tetartan (blue-yellow defects due to post receptor loss) colour defects.[4] The HRR test has 24 plates each displaying either one or two figures of a square, circle, or triangle. The first four plates (1-4) are for demonstration, and the next six (5-10) are for screening. The first two screening plates (5-6) each have two figures which are designed to screen for blue-yellow colour-vision deficiencies. The next four plates (7-10) present a total of six figures designed to screen for red-green colour-vision deficiencies. The last 14 plates (11-24) attempt to classify the nature of the defect and estimate the severity based on the type and number of errors made.[8]

All plates are the vanishing design. The background dots are grey, and the dots that make up the geometric symbol are printed in the colours which are confused with grey for protan, deutan, and tritan colour deficiency. If the figures on plates 5-10 are seen correctly, then the subject is considered to have normal colour vision and test is stopped.[8] The value of the HRR test is in classifying protan and deutan defects, grading the severity of red-green colour deficiencies and identifying moderate tritans.[4]

The HRR (third edition, Richmond Products, 1991) was used in this study. This test attempted to reproduce the first two editions, but the manufacturer fell short in three areas. First, the colours are poorly aligned with respect to the deuteranopic lines of confusion which can result in deutans passing the test or underestimating the severity of the defect. Second, it is less saturated than original AO-HRR and may overestimate the degree of defects. Third, it was reported that the colours are a metameric match and not a spectral match and so there may be different error rates with different light sources.[9] Although the third edition is not as good as the original based on the colorimetric analysis, it was included because no information was available on its clinical performance and the fourth edition wasn't available at the time.

In addition to obtaining information on its clinical utility, the 3rd Edition of the HRR plates was included because previous studies showed that the 1st and 2nd Editions could predict performance in the ability of colour defectives to recognize colour codes of resistors and these results suggested the test may be useful in predicting performance in naming VDT colours.[10]

3.1.4. Farnsworth D15 test

The Farnsworth Dichotomous D-15 (D-15) test is used to separate colour-defective individuals who could perform adequately on most daily colour-related tasks from those who would likely encounter difficulties.[11] In general, the D-15 succeeds in meeting this goal and so it is often recommended as a test used to assess occupational fitness of colour defective individuals including police and firefighters. [6, 11-13]

The Farnsworth D15 contains 16 coloured caps. Each cap has a different Munsell colour which was selected so that the difference in hue between adjacent colours was approximately equal around the colour circle. The colour differences between adjacent numbered caps are sufficiently large so that colour vision normals and mild colour-defectives can complete the test without error. One cap is fixed as a reference and the other 15 are moveable. The 15 moveable caps all have a Munsell value (luminous reflectance) of 5 and a chroma (saturation) of 4. The colour caps subtend 1.5 degree at 50cm. The caps are numbered on the back.

The moveable colour caps are randomly arranged in front of the subject. There are two different sets of instructions that can be used. One would be to instruct subjects to arrange the colour caps in the order of colour starting at blue and ending in purple. The other way is to instruct them to arrange the caps in the box by finding the moveable cap that is most similar in colour to the last one place in the box. The second way does not require subjects to have a concept of colour order.

The test results can be evaluated by visual inspection or numerical analysis such as the sum of the colour differences between adjacent caps or vector analysis.[14, 15] For visual inspection, the standard score sheet is used. The score sheet has the numbered caps arranged in a circular pattern to represent the hue circle. The numbers are connected with lines drawn according the subject's arrangement. Major errors of arrangement produce lines which cross either the horizontal or vertical meridian of the circle. A major crossing is also defined as a difference between adjacent caps which is greater than 2. Two or more crossings are generally regarded as a failure and indicated the presence of a more severe colour vision defect.[11]

3.1.5. Adams desaturated D15 test

The Adams desaturated D-15 (Adams D-15) colour vision test was introduced to provide a test for acquired colour vision defects that was more sensitive than the standard Farnsworth D-15 and quicker to administer than the Farnsworth-Munsell 100 hue test.[16,17] The Adams D-15 caps have the same hue and value as the standard D-15 but the chroma level is reduced by two units to a chroma of two.[16]

The Adams D-15 can also be useful as part of a test battery for congenital defects. Dain and Adams have shown that the Adams D-15 can be used to grade the severity of the defect into mild (pass the Adams D-15 and standard D-15), moderate (fail the Adams D-15 but pass the standard D-15) and severe (fail both tests).[18] A recent study by Hovis et al[17] shows that the test also has reasonable repeatability for failure criteria of any one major crossings and more than one major crossing. The test procedure is identical to the D15 test.

The Adams D-15 was chosen for this study because of its reported increased sensitivity relative to the Farnsworth D-15 and we were curious as to whether the Adams D-15 can predict performance of colour defectives on the VDT task.

3.1.6. CN Lantern test

This CN Lantern test provides a reasonable functional assessment of one's ability to identify wayside rail signal light colours.[19] The CN Lantern presents 15 different triplets of test lights. The first two triplets are demonstration and the next 13 are the test sequence. The colours displayed on a given trial could be any combination of red, green, and yellow lights. Each light subtends a visual angle of 1.25 arc min at 4.6 m. This angular size is equivalent to a sighting distance of 0.5 km of the actual wayside signals. The intensities of the lights vary between 6 to 22 times greater than their detection threshold for the background luminance.[19]

The testing protocol starts with the subject seated at a 4.6 m viewing distance. If the person fails the test, it is repeated at 2.3 m. The different viewing distances are used to represent two different viewing distances encountered in the railway industry. The 4.6 m is used to test individuals who may be working on the main track where the sighting distances for signal lights are usually near 0.5 km and the shorter distance is used to test individuals who may be working in the rail yard where the sighting distances for signal lights are typically less than 0.25 km.

The response to each test light is scored as either correct or incorrect. A single error is allowed at the 4.6 m test distance provided that the mistake was not identifying a red light as green or vice versa. No errors were allowed at the shorter test distance. The different pass/fail criterion for the two test distances was based on the result that the colour-defectives who made errors at 4.6m had decreased at a shorter distance.[20]

Although CN Lantern is not a clinical colour vision test, it was included in the test battery because it was already being used by the railways and could be a useful predictor of performance on the VDT based test. More importantly, we wish to determine whether a person who passed the CN Lantern can pass the VDT test. If this was the case, then employees who pass the CN lantern test may not have to take the VDT test.

References

1. Hardy, L., G. Rand, et al. (1945). "Test for the detection and analysis of colour-blindness: I. The Ishihara test: An evaluation." J Opt Soc Am A 35: 268-275.
2. Birch, J. and L. M. McKeever (1993). "Survey of the accuracy of new pseudoisochromatic plates." Ophthalmic Physiol Opt 13(1): 35-40.
3. Birch, J. (1997). "Efficiency of the Ishihara test for identifying red-green colour deficiency." Ophthalmic Physiol Opt 17(5): 403-8.
4. Birch, J. (2001). "Diagnosis of defective colour vision, Second edition." Chapter 9: 108-109.
5. Pease, P. L. and J. Allen (1988). "A new test for screening colour vision: concurrent validity and utility." Am J Optom Physiol Opt 65(9): 729-38.
6. CIE Technical Report (2001). "International recommendations for colour vision requirements for transport.13-14"
7. Pokorny J, S. V., Verriest G, Pinckers AJLG eds (1979). "Congenital and Acquired Colour Vision Defects." Newyork: Grune and Stratton.
8. Birch, J. (1997). "Clinical use of the American Optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates for red-green colour deficiency." Ophthalmic Physiol Opt 17(3): 248-54.
9. Dain, S. J. (2004). "Colourimetric analysis of four editions of the Hardy-Rand-Rittler pseudoisochromatic tests." Vis Neurosci. 21(3): 437-43.

10. Walraven, P. and H. Leebeck (1960)."Recognition of colour code by normals and colour defectives at several illumination levels. An evaluation study of the H.R.R. plates." Am J Ophthalmol(Feb;37:): 82-92.
11. Farnsworth, D. (1947). "The Farnsworth Dichotomous Test for Colour Blindness Panel D-15. New York: The Psychological Corp."
12. Margrain TH, Birch J, et al. (1996). "Colour Vision Requirements of Firefighters." Occupational Medicine 46: 1996;46:114-124.
13. Birch, J. (2008). "Pass rates for the Farnsworth D15 colour vision test." Ophthalmic Physiol Opt. 28(3): 259-64.
14. Bowman, K. J. (1982). "A method for quantitative scoring of the Farnsworth Panel D-15." Acta Ophthalmol (Copenh) 60(6): 907-16.
15. Vingrys, A. J. and P. E. King-Smith (1988). "A quantitative scoring technique for panel tests of colour vision." Invest Ophthalmol Vis Sci 29(1): 50-63.
16. Adams AJ, R. R. (1982). "Use of desaturated and saturated versions of the D-15 test in glaucoma and glaucoma-suspect patients." In:Verriest G, editor. Colour Vision Deficiencies IV. Doc Ophthalmol Proc Series 33.The Hague: Dr W Junk; 33: 419-424.
17. Hovis, J. K., S. Ramaswamy, et al. (2004). "Repeatability indices for the Adams D-15 test for colour-normal and colour-defective adults." Clin Exp Optom 87(4-5): 326-33.
18. Dain, S. J. and A. J. Adams (1990)."Comparison of the standard and Adams desaturated D-15 tests with congenital colour vision deficiencies". Ophthalmic Physiol Opt 10(1):40-5.

19. Hovis, J. K. and D. Oliphant (2000). "A lantern colour vision test for the rail industry." Am J Ind Med. 38(6): 681-96.
20. Hovis, J. K. and S. Ramaswamy (2006). "The effect of test distance on the CN lantern results." Vis Neurosci. **23**(3-4): 675-9.

CHAPTER 4

Methods

4.1. Subjects

One hundred subjects with normal colour vision and 52 subjects with congenital red-green defects participated in the study. The subjects were recruited through newspaper advertisements, electronic bulletin boards, and posters. Table 4.1 shows the prevalence of the different types of defects in the colour-defective group. There was a significantly higher proportion of protanamolous and dichromatic defects and a corresponding fewer proportion of deuteranamalous individuals than expected based on a Caucasian population (X^2; $p > 0.05$ rejection level). This difference may have been due to the recruiting process which advertized for subjects in newspapers, flyers and electronic bulletin boards (Appendix 4.1).

Table 4.1. Prevalence of the different types of red-green defects in the sample and the expected prevalence in Caucasian population for comparison.

Type of Defect	Percent in First Session (N = 52)	Percent in Second Session (N = 44)	Percent in Caucasian Colour-Defective Population
Deuteranomalous	52	55	64
Protanomalous	30	27	12
Deuteranope	8	7	12
Protanope	10	11	12

All the subjects had a minimum binocular visual acuity of 6/9 for distance and near (i.e. 40 cm). The mean age of the colour-normals was 29 years (SD ± 9.7) and the mean age of the colour-defectives was 30 years (SD ± 13). These means were not significantly different (t-test; $p<0.05$ rejection level). Each subject was invited to return in approximately 10 days time to repeat the test. Ninety three colour-normals (93%) and 44 (85%) colour-defectives participated in both sessions. The prevalences of the different red-green defects were statistically similar to the first session (X^2; $p \geq 0.05$).

4.2. Clinical Colour Vision Test Procedures

4.2.1. Ishihara Procedure

The subjects were separated as colour-normals and colour-defectives using the 38 plate edition of the Ishihara test. The test was administered at viewing distance between 40 and 60 cm distance under an Illuminant C light source. We used the current Railway Association of Canada's screening standard of 6 or more errors on the screening plates (plates 2-21) as a failure. Birch reports that this criterion gives a sensitivity of 98% and specificity of 100%. [1]

4.2.2. Nagel Anomaloscope Procedure

The Nagel anomaloscope in the white adaptation mode was used to confirm the Ishihara's screening results and further classify the defect. The Nagel anomaloscope testing procedure varied depending on whether the Ishihara classified the subject as colour-normal or colour-deficient. The procedure was as follows. For subjects with normal colour vision, a red-green mixture was set between 35-40 units on red-green scale and the subject was asked to make a hue match by adjusting the red-green mixture knob. Then they are asked if the top and bottom halves are of same brightness. If not, they are asked to make a brightness match and the values are noted. This gives the midpoint or matching ratio. This procedure was repeated three to four times.

Next the range of acceptable matches was determined. The experimenter adjusted the red-green either + 5 units from the median match point setting. Next the subject made a brightness match by adjusting the yellow intensity. They were then asked whether the two fields were identical in colour and brightness. If there was not a brightness match, the experimenter changed to the red-green setting to +2 units and the process was repeated in progressively smaller steps until the range on that side of the midpoint was determined. Next the range of acceptable matches was determined for the opposite direction from the median match using the same bracketing procedure.

In the case of colour-defective subjects, the red-green scale was set at either 0 (maximum green) or 72 (maximum red), Subjects were asked to make a brightness match between the two halves by adjusting the intensity of the yellow field. Once the brightness match was obtained, then the subject was asked whether the two fields looked identical in colour. Their responses and yellow intensity setting was recorded and the red-setting was adjusted by 10 and the process was repeated until the entire range of the red-green settings was assessed. After completing this part of the assessment, the responses were reviewed by the examiner. If there were colour matches during the initial phase, the bracketing procedure was used except the match settings were used as the starting value instead of the midpoint. If no matches were present, then the procedure used for colour-normals was implemented. Consistent with the white adaptation mode, all subjects were instructed to look at the white screen approximately every 10 sec for approximately 2 sec throughout their settings.

In addition to the Ishihara test and the anomaloscope, colour vision was also assessed with the HRR pseudoisochromatic plates (Third edition, Richmond products, 1991) Farnsworth D-15, Adams Desaturated D-15 and the CN Lantern test. The order of testing was Ishihara test, Nagel anomaloscope, HRR pseudisochromatic plates, Farnsworth D-15, Adams Desaturated D-15 and the CN Lantern test.

4.2.3. HRR Pseudo-Isochromatic Plate Test Procedure

The HRR (third edition, Richmond Products, 1991) was administered at viewing distance between 40 and 60 cm distance under an Illuminant C light source. If all the figures on plates 5-10 were seen correctly, then the subject was considered to have normal colour vision and test was stopped.[2] If the subject made an error, s/he was tested with the diagnostic plates (11-24) to assess the type and severity of the defect. The total errors made on the diagnostic plates were recorded. If the subject did not make any errors on the diagnostic plates, then the screening plates was redone. The severity is usually based on the errors made on the protan and deutan diagnostic plates. In this study, we added the errors in both the protan and deutan diagnostic plates. Given some of the limitations reported for the 3rd edition,[3] we felt this protocol would reduce the potential of underestimating the severity of the deutan defects and we felt that this would potentially classify individuals with enlarged discrimination ellipses in both directions as more severe. The severity was classified using our procedure very mild (0-1 error), mild (2-3 errors), moderate (4-5 errors) and severe (6-8 errors) based on the number of errors in the diagnostic plates. In this study, for all analysis purposes the total number of errors on the red-green diagnostic plates was considered. Note that none of the colour-defectives in this study performed a blue-yellow error.

4.2.4. Farnsworth and Adams Desaturated D-15 Procedures

The D15 tests contain 16 coloured caps. One cap is fixed as a reference and the other 15 are moveable. The moveable colour caps are randomly arranged in front of the subject. Subjects were instructed to arrange the caps in the box by finding the moveable cap that is most similar in colour to the last one place in the box. The subjects were not timed and were allowed to take as long as is necessary to complete the test.

The results were recorded on the standard score sheet and evaluated by visual inspection. The score sheet has the numbered caps arranged in a circular pattern to represent the hue circle. The numbers are connected with lines drawn according the subject's arrangement. Major errors of arrangement produce lines which cross either the horizontal or vertical meridian of the circle. Two or more, major crossings were a failure.

4.2.5. CN Lantern Test Procedure

The testing protocol starts with the subject seated at a 4.6 m viewing distance under room illumination of 300 lux in a plane, parallel to the floor and at the height of the each test. The subjects were informed that three different lights would be displayed and they had to identify the colour of each light. The colours could be any combination of red, green or yellow. Examples of each colour were shown once before the start of the actual testing.

The response to each test light is scored as either correct or incorrect. A single error is allowed at the 4.6 m test distance provided that the mistake was not identifying a red light as green or vice versa.[4] If the person fails the test, it is repeated at 2.3 m and then at 1.15 metres. No errors were allowed at the shorter test distance.

4.3. Test Procedure for the Practical Test

The test procedure details are described in Appendix 2.1. A brief summary is presented below. The subject was seated at one metre away from the test monitor. This is the usual viewing distance in the network centre. Room illumination was dim at 75 lux in the plane of the table where the monitor was placed. The name and identification number of the employee (or subject) was entered in the first screen displayed. The instructions for each section were similar in that the subjects were told to identify the colour of each icon by clicking on the response circle next to the appropriate colour name (or just click on the appropriate name) name using a mouse. Each section is preceded by more specific instructions and examples of the colours. The subject could change their responses any time on a given screen, but once they clicked the "proceed to the next section", their responses were saved and they could not revert back to the previous screen. They were naive as to the number of times each colour was presented on a screen. The subjects were allowed to proceed through the test at their own pace. Although there was no time limit imposed for completing the test, the time required to perform the test was recorded for most individuals.

The University of Waterloo's Office of Research Ethics reviewed the study to ensure that it met the ethical guidelines. Subjects who participated in the study gave informed written consent before participating.

The next chapter presents the validation results of the practical test.

References

1. Birch, J. (2001). "Diagnosis of defective colour vision, Second edition." Chapter 9: 108-109.
2. Birch, J. (1997). "Clinical use of the American Optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates for red-green colour deficiency." Ophthalmic Physiol Opt 17(3): 248-54.
3. Dain, S. J. (2004). "Colourimetric analysis of four editions of the Hardy-Rand-Rittler pseudoisochromatic tests." Vis Neurosci. 21(3): 437-43.
4. Hovis, J. K. and S. Ramaswamy (2006). "The effect of test distance on the CN Lantern results." Vis Neurosci. 23(3-4): 675-9.

CHAPTER 5

Results I. VDT Test First & Second Session Results and Repeatability

5.1. VDT test –First session

Figure 5.1 shows the distribution of total errors made by colour-normals and colour-defectives in the first session of the VDT test. The figure shows that the colour-normals make very few errors with a mode of zero, a mean of 0.30 and a maximum value of 6. The performance of the colour defectives was more varied with a mean of 7.8 and a maximum score of 29. The mode for the colour-defective group was also zero, although the frequency of the number of individuals who performed perfect was only 28 % compare to 84 % for the colour-normal group.

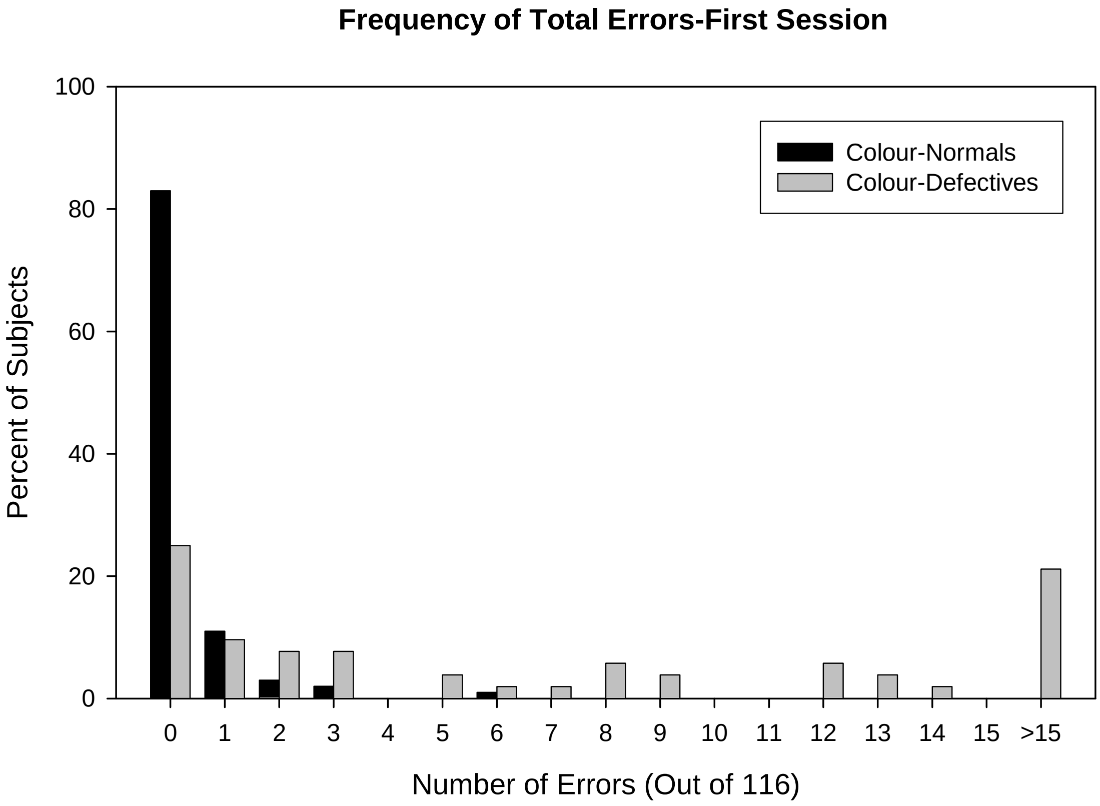


Figure 5.1 Distribution of total errors for colour-normals and colour-defectives in the first session of the VDT test.

With respect to the individual sections of the test, no one made any errors on the signal icon test colours. This finding indicates that the colour-defectives are able to correctly identify the red and green icons when there are the only these two possible colours.

Figure 5.2 and 5.3 show the frequency of errors made by colour-normals and colour-defectives in the track grid (second section of the test) and occupancy authority (third section of the test) parts of the VDT test for the first session. The frequency of errors on the track grid and occupancy authority shows the same trends as the total number of errors; namely colour-normals made few errors and colour-defectives had a varied performance. Zero was again the mode of the distributions for both groups and sections.

The major difference between the two sections was that errors were more frequent on the authority colours section of the tests for both the groups. The mean error for the colour-normals was 0.06 on the track grid section and 0.24 on the authority section. The colour-defectives means for the track grid and authority sections were 3.4 and 4.4 respectively; however, the means for the respective groups were not significantly different (Paired t-test: $p \geq 0.05$). Since some colour-normals did make errors and several colour-defectives had performances that were within the normal range, it is important to see what types of errors were made by the colour-normals and these colour-defective subjects.

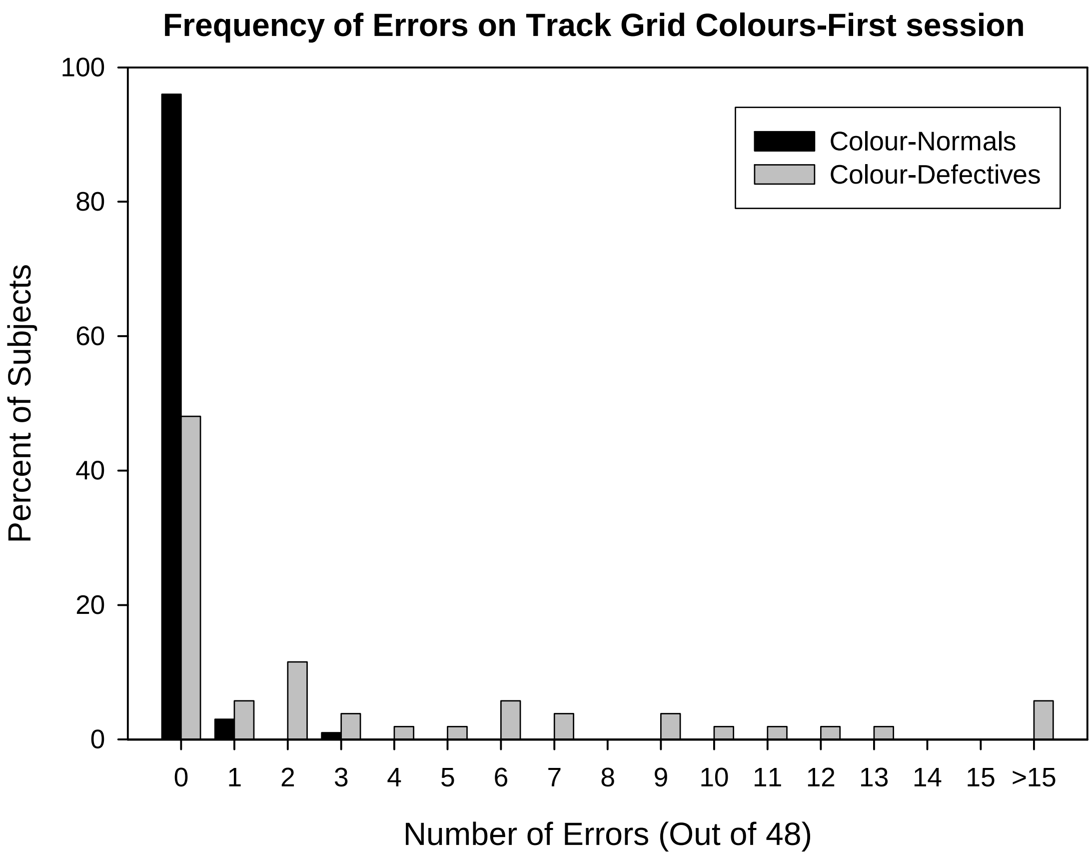


Figure 5.2 Distribution of errors made by colour-normals and colour-defectives in the track grid (second) part of the VDT test in the first session.

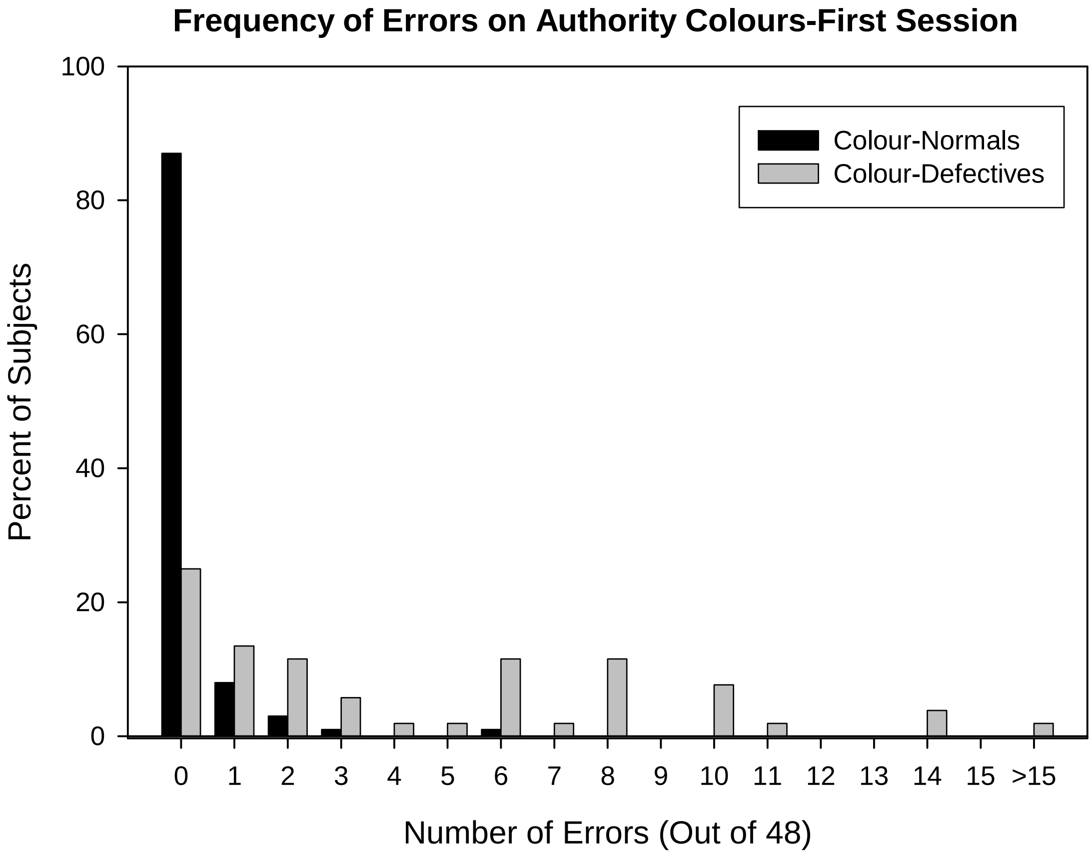


Figure 5.3. Distribution of errors made by colour-normals and colour- defectives in the authority (third) part of the VDT test in the first session.

Tables 5.1 a, b, c & d shows the types of errors made by both groups in the track grid and occupancy authority sections of the test. The first part (signal icon) of the VDT test was omitted since none of the subjects in either group made any errors. The percentages given are relative to the total number of errors made by all subjects within each group. With this analysis, a relatively high percentage would result if either the error was common across several subjects or a few subjects made the same error multiple times.

Table 5.1a.Types of errors made by colour-normals on the track grid section of the test during the first session. Percentages are based on the total number of errors made by colour-normals on this section which was 6. Shaded cells highlight errors that are more frequent. These errors are consistent with tritanopic lines of confusion.

Response>> **Colour**	**Red**	**Yellow**	**Green**	**White**	**Grey**	**Blue**	**Purple**	**Blue Green**
Red		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Yellow	0.0%		0.0%	0.0%	16.67%	0.0%	0.0%	0.0%
Green	0.0%	0.0%		0.0%	16.67%	0.0%	0.0%	0.0%
White	0.0%	16.67%	0.0%		0.0%	0.0%	0.0%	0.0%
Grey	0.0%	0.0%	0.0%	0.0%		0.0%	16.67%	0.0%
Blue	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%	0.0%
Purple	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%
Blue Green	0.0%	0.0%	0.0%	0.0%	0.0%	33.33%	0.0%	

Table 5.1b. Types of errors made by colour-defectives on the track grid section of the test during the first session. Percentages are based on the total number of errors made by colour-normals on this section which was 179. Shaded cells highlight errors that are consistent with either the protanopic or deuteranopic lines of confusion.

Response>> **Colour**	**Red**	**Yellow**	**Green**	**White**	**Grey**	**Blue**	**Purple**	**Blue Green**
Red		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Yellow	0.0%		10.06%	0.0%	0.0%	0.0%	0.0%	0.0%
Green	0.0%	31.84%		0.0%	0.0%	0.0%	0.56%	0.0%
White	0.0%	0.56%	0.0%		0.0%	0.0%	0.0%	0.0%
Grey	1.12%	0.0%	0.56%	0.56%		0.0%	4.47%	8.94%
Blue	0.0%	0.0%	0.0%	0.0%	0.0%		3.91%	0.0%
Purple	0.0%	0.0%	0.0%	0.0%	0.0%	11.73%		2.79%
Blue Green	0.0%	0.0%	2.79%	1.12%	12.29%	2.23%	4.47%	

Table 5.1c. Types of errors made by colour-normals on the Track section of the test during the first session. Percentages are based on the total number of errors made by colour-normals on this section which was 25. Shaded cells highlight errors that are more frequent. These errors are consistent with tritanopic lines of confusion.

Response>> Colour	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Orange	44.0%		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Yellow	0.0%	4.0%		0.0%	0.0%	0.0%	0.0%	0.0%
Green	0.0%	0.0%	0.0%		0.0%	0.0%	0.0%	12.0%
Dark green	0.0%	0.0%	0.0%	12.0%		0.0%	0.0%	0.0%
Grey	0.0%	0.0%	0.0%	4.0%	0.0%		0.0%	0.0%
Blue	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		8.0%
Light Blue	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.0%	

Table 5.1d. Types of errors made by colour-defectives on the occupancy authority section of the test during the first session. Percentages are based on the total number of errors made by colour-normals on this section which was 224. Shaded cells highlight errors that are consistent with either the protanopic or deuteranopic lines of confusion.

Response>> Colour	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red		2.68%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Orange	28.57%		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Yellow	0.0%	0.45%		11.16%	0.0%	0.0%	0.0%	0.0%
Green	0.0%	2.68%	29.46%		0.45%	0.0%	0.0%	0.0%
Dark green	0.0%	0.0%	0.0%	8.48%		8.04%	0.0%	0.0%
Grey	0.0%	0.0%	0.0%	0.45%	7.14%		0.0%	0.45%
Blue	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%		0.0%
Light Blue	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

5.1.1. Colour-Normal Errors

Tables 5.1a and c show the types errors made by colour-normal’s in the track (second part) and occupancy authority (third part) of the VDT Dispatch test. The tables show that the most common error in colour-normal’s was misnaming orange as red. The relatively high percentage of the orange-red mistakes was due to the fact that 4% of colour normal subjects made this mistake at least once and half of these individuals made this error several times within the session. That is, 2% of the colour-normal group was consistent in making this mistake through the test. The orange-red error was not surprising given that the small difference in colour and luminance between the red and orange (Chapter 2. Figures 2.3 and 2.4).

Interestingly, this error and the ones that are shaded in Tables 5.1a and 5.1c are consistent with a tritan defect if luminance differences are ignored and it may be possible that some colour-normal individuals made tritan errors (irrespective of age) because they weren’t viewing the stimuli for a sufficient duration or weren’t centrally fixating on the icon or their colour. More likely, however, was that they had relatively poorer colour discrimination which resulted in large discrimination ellipses in these regions of the colour space. Most of the discrimination ellipses measured by MacAdam have their long axis oriented along the tritanopic confusion lines and so a few tritan-like errors are not surprising. Nevertheless, this does raise the issue as to whether subtle tritan defects were missed. Although one cannot rule this out completely, none of these subjects made any errors on the HRR blue-yellow plates which have colour differences smaller than the ones displayed on the VDT test. It seems unlikely that we missed any clinically significant tritan defects which would translate into numerous errors on the VDT test.

The other errors made by colour-normals are more difficult to explain. One possibility is that these errors were due to a data entry. For example, the “blue-green called blue” error may have occurred because the subject mistakenly clicked on the circle for the blue response which is just below the blue-green in the list of possible answers (See Figure 2.7 in Chapter 2). These types of data entry errors may be more likely for responses in listed in the middle of the response column, such as green, dark green, blue and light blue, and this may explain why errors occurred frequently on these colours.

5.1.2. Colour-Defective Errors

As expected, Tables 5.1 b and 5.1 d show that the colour-defectives made the most errors on colours which are near the red-green lines of confusion ignoring any luminance differences. Their most frequent single type of error was yellow misnamed as green. If the yellow misnamed as green and green misnamed as yellow errors are pooled, then confusing yellow and green colours with each other was the most frequent pair of colours confused (40% of the total errors). Yellow-green confusions are typically made by individuals with red-green colour vision defects and the luminances of these two colours were similar. Nevertheless, it is possible, that the mistakes could be a data entry error if it occurred in the last section of the test because the green response was just below the yellow response in the authority colour section. A data entry error, however, is unlikely since none of the colour-normals made a yellow-green error. The more probable reason would be that the error was a result of the underlying colour vision problem.

The types of errors made by colour-defectives will be discussed in more detail in the next chapter.

5.2. VDT test –Second session

Figure 5.4 shows the frequency of the total errors made by the colour-normal and colour-defective groups at the second session. Figure 5.5 and 5.6 shows the frequency of errors scores for the track grid (second part) and occupancy authority (third part) colour sections. In general, the results of the second session were similar to the first session. Colour-normals made few errors, the colour-defective results were more variable, and the errors on the occupancy authority colours were the most frequent for both groups. Comparing Figures 5.4, 5.5 and 5.6 with Figures 5.1, 5.2 and 5.3 one notices that there was an improvement in performance at the second session. In particular, the percentage of subjects in both groups with perfect scores was higher in the second session. The improved performance was also reflected in the mean number of errors. The mean total error for the colour-normal group at the second session was 0.09 with a maximum error score of 2, whereas the mean for the colour-defective group was 4.8 with a maximum error score of 17. In the track grid section, the mean error score for the colour-normals and colour defectives was 0 and 1.91 respectively, whereas the mean number on the authority section was 0.08 and 2.86 for the colour-normals and the colour-defectives respectively.

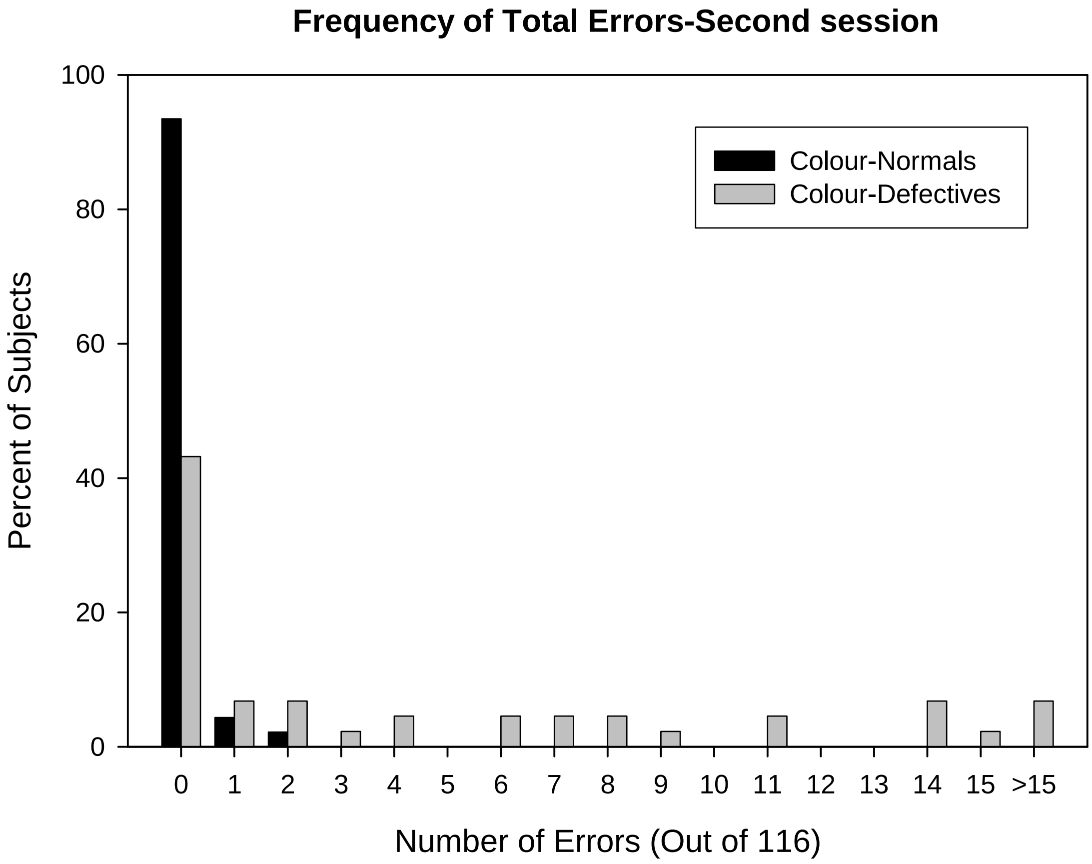


Figure 5.4. Distribution of total errors made by colour-normals and colour-defectives at the second session of the VDT test.

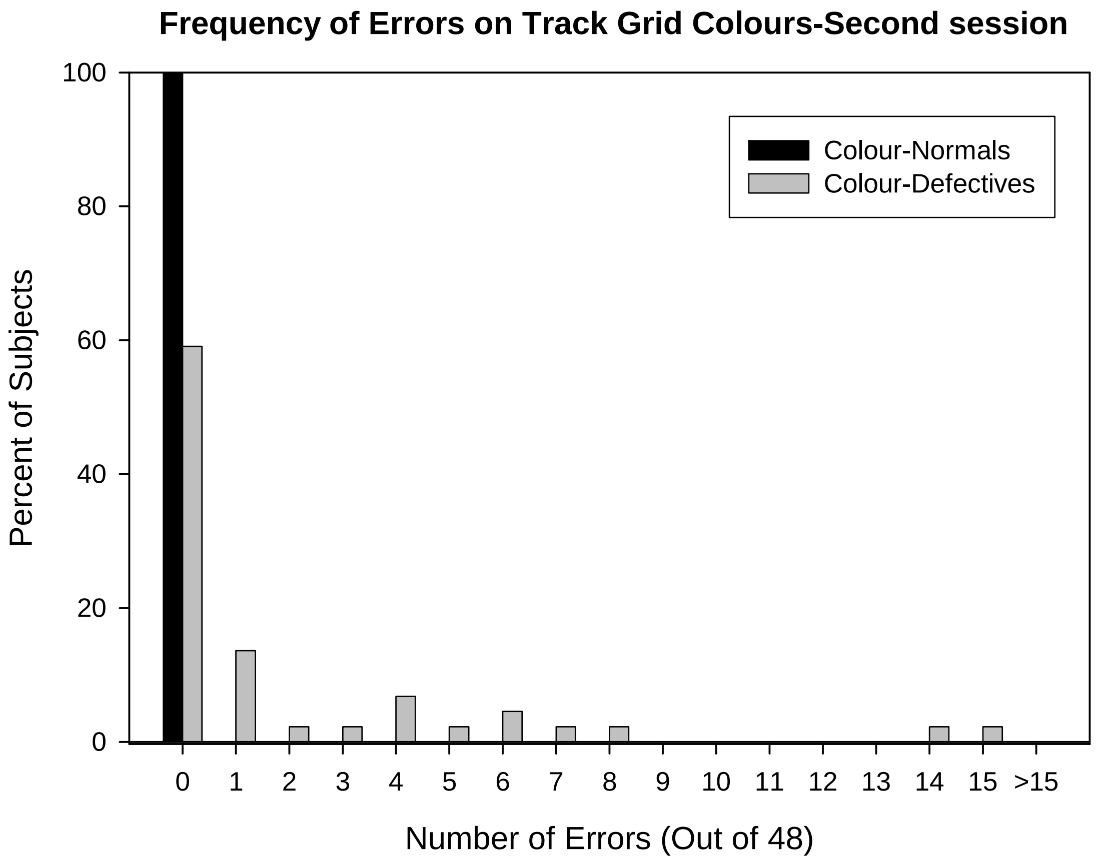


Figure 5.5. Distribution of errors made by colour-normals and colour-defectives in the track grid part of the VDT test at the second session.

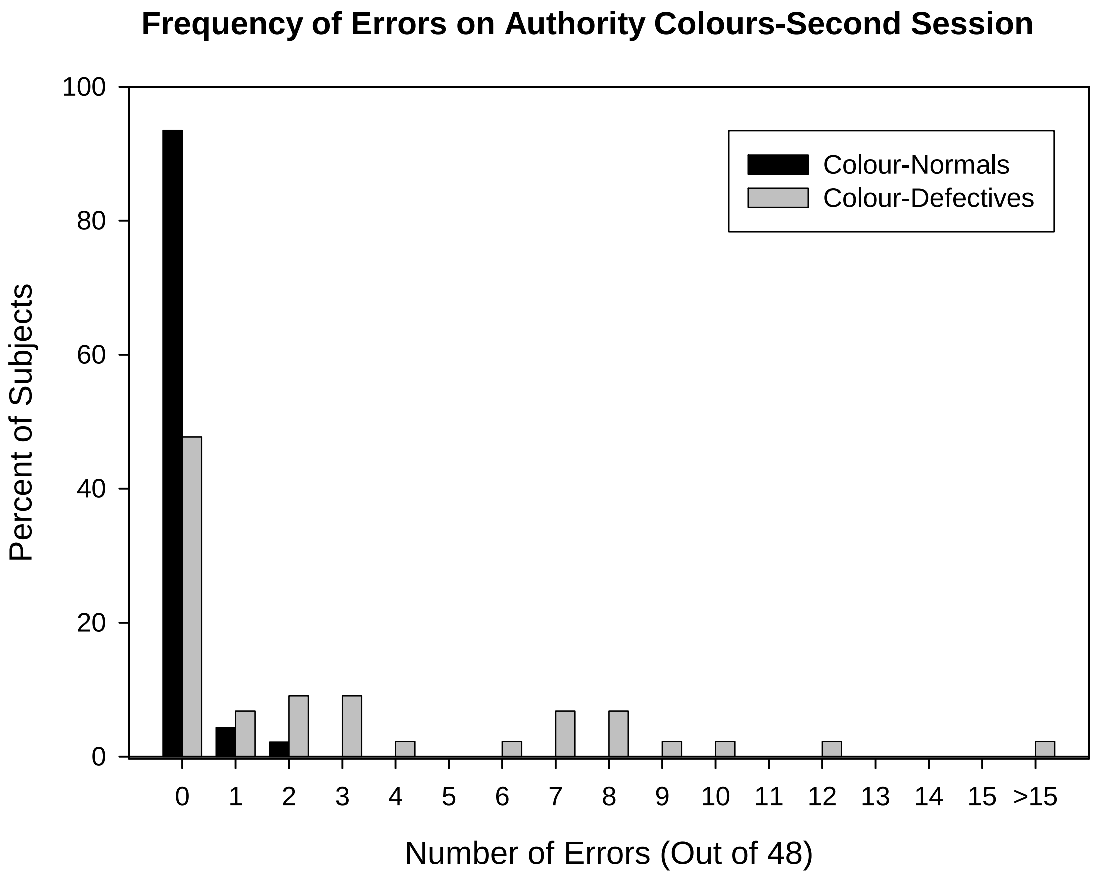


Figure 5.6 Distribution of errors made by colour-normals and colour-defectives in the authority section of the VDT test at the second session.

Table 5.2. Mean difference in the total error score of subjects who participated in sessions. Negative values indicate there was an improvement in the error score at the second session.

Subject Group	**Mean Errors-First session**	**Mean Errors-Second session**	**Mean Difference (Second –First)**	**Standard deviation**	**95% Confidence Interval**
Colour-Normals	0.29	0.09	-0.20	1.00	-0.03 to -0.43
Colour-Defectives (CVD)	6.98	4.8	-2.18	4.82	-3.60 to -0.74

The improved performance at the second session could be due to attrition by the worst-performing subjects, learning/practice effects, or both. To control for the attrition effect, the mean difference in the total error scores for those who participated in both sessions were calculated and are shown in the Table 5.2. The table shows that the mean difference was significantly less than zero (based on the 95% confidence intervals) for both groups indicating that there was a significant learning/practice effect, although it was small in terms of the absolute change.

The types of error and their relative frequencies showed a similar trend as the first session. The most common type of error for the colour-normals was misnaming orange as red, whereas the errors made by colour-defectives were most frequent for colours that lie near the lines of confusion and, again, yellow-green confusions were the most common mistake. This type of performance was exhibited by a small percentage (4%) of the colour-normal's who misnamed orange as red at the first session, but identified it correctly on the second session. There also were two colour-normal individuals (2%) who made this error only on the second trial.

5.3. Establishing a Pass/fail score for the VDT test and Repeatability

Figures 5.1 and 5.4 showed that approximately 40 % of the colour-defectives performed as well as colour-normals in identifying colours. There were, however, some colour-defectives who made a small number of errors that were more consistent with red-green colour vision defect rather than normal colour vision. Although the number of errors was small, they could have important safety consequences. There was also a significant learning effect indicated by the decrease in the number of errors on the second session. Together, these findings suggest that the pass/fail criterion may have to be based on the number and types of errors along with a consideration of whether the person is repeating the test.

The intent of the VDT test is that it will be administered to anyone applying for a dispatcher's position who fails the colour vision screening test. Individuals who pass the screening test will not have to undergo subsequent testing. This intention is consistent with the employment practices of excluding anyone who fails the colour vision screening from the dispatcher position, but allow everyone who passes the screening test to apply. One could argue that given the safety implications all candidates should take the practical based test; however, other than ensuring that the orange and red occupancy authority colours are sufficiently different, the results of this study do not make a strong case for testing everyone who is applying for the dispatch position with the VDT test.

Given that the colour-normals will not be taking the test, the purpose of establishing an appropriate pass/fail score is to find a maximum error score for which the majority of colour-normals would pass, the repeatability is high and safety is not compromised. Ideally we would want all colour-normals to pass, as they would not be administered the practical test once they pass the screening test.

A number of scoring criteria where both the number and types of errors determine the pass/fail score were considered. Allowable errors were based on the errors made by colour-normals listed in Tables 5.1a and 5.1c that would not be near the red-green dichromatic lines of confusion. The one exception was the orange-red mistake because this was the most frequent colour-normal error. Table 5.3 lists these errors. The learning effect is also taken into account in many of these criteria by reducing the number of allowable errors on the second session. Table 5.4 lists the scoring criterion evaluated and their basis.

Table 5.3. Mistakes that could be allowed on the VDT test based on the colour-normal responses.

Test Colour	**Acceptable Mistake**
Orange	Red
Yellow	Grey or White
Green	Light blue or Grey
Blue-green	Blue or Green
White	Yellow
Grey	Green
Blue	Light blue or Grey

Table 5.4. Possible scoring criteria for the VDT test and their basis. (The percentile scores are rounded to the nearest integer value).

Criterion Label	**Definition**	**Basis**
1	Allow only 1 error on either trial	95th percentile score of colour-normals from first trial
1*	Allow only 1 error on either trial, but only errors listed in Table 5.3 are permitted.	95th percentile score of colour-normals from first trial
2	Allow 2 errors on either trial	97th percentile score of colour-normals from the first trial
2*	Allow 2 errors on either trial, but only errors listed in Table 5.3 are permitted.	97th percentile score of colour-normals from the first trial
3	Allow 3 errors on either trial	Average error score of worst-performing colour-normal (rounded down to the nearest integer)
3*	Allow 3 errors on either trial, but only errors	Average error score of worst-performing

	listed in Table 5.3 are permitted.	colour-normal (rounded down to the nearest integer)
1/0	Allow one error on first trial and none on second	95th percentile scores of colour-normals on each session
1/0*	Allow one error on first trial and none on second, but only errors listed in Table 5.3 are permitted.	95th percentile scores of colour-normals on each session
2/1	Allow 2 errors on first trial and 1 on the second	97th percentile score of colour-normals on each session
2/1*	Allow 2 errors on first trial and 1 on the second, but only errors listed in Table 5.3 are permitted.	97th percentile score of colour-normals on each session
2/1**	Ignore orange-red errors on both trials, and then allow 2 errors on first trial and 1 error on the second.	98th percentile score on the revised error score on first session and the worst-normal on the second session
4/1	Allow 4 errors on first trial and 1 on second	99th percentile score of colour-normals on each session
4/1*	Allow 4 errors on first trial and 1 on second but only errors listed in Table 5.3 are permitted.	99th percentile score of colour-normals on each session
7/2	Allow 7 errors on first trial and 2 errors on the second	Worst colour-normal on each session
7/2*	Allow 7 errors on first trial and 2 errors on the second, but only errors listed in Table 5.3 are permitted.	Worst colour-normal on each session

5.3.1. Repeatability

To determine which criterion had the best repeatability, several different measures were used. The primary index was the κ (kappa) coefficient of agreement for repeatability for the colour-defective group. This value measures the between-session agreement after correcting for chance agreements. A κ value of 0 indicates no agreement beyond chance, whereas a kappa value of 1 indicates perfect agreement.[1]

The κ coefficient was calculated for only the colour-defective group for three reasons. First, the test was designed to be administered to only colour-defective employees or perspective employees and so there is a need to establish the repeatability for that particular group. Second, including the colour-normal with the colour defective subjects could artificially give a high κ value because most colour-normals passed at both sessions. Third, calculating a separate κ coefficient for the colour-normals would be meaningless since only few colour-normals failed on the both sessions.[2]

The second and third indices for repeatability were the proportions of the two types of between-session discrepancies that could occurred. One type of between-session discrepancy was the proportion of colour-defectives who passed the first session, but then failed the second session. The other discrepancy was the proportion that failed on the first session, but passed on the second.

A criterion which exhibits an extremely high κ coefficient will also have the lowest proportion of discrepancies; however, analyzing the proportions of the different types of discrepancies between sessions and the proportion of colour-normals who pass both sessions would help decide between the various options when the κ coefficients are relatively high and nearly equal to one another. The fourth index of repeatability is the proportion of colour-normals who pass on both trials. This value is calculated to ensure that the proportion of colour-normals who pass at both sessions is high.

The κ coefficients and different proportions are shown in Figures 5.7 through 5.9. Figure 5.7 shows that, barring one exception (the criterion 2/1**), making the pass/fail score contingent on the types of errors lowers the repeatability of the test for colour-defectives. The single exception of criterion 2/1** was not counting misnaming orange as red as an error.

There were two reasons for the lower repeatability seen for most scoring criterion contingent on the types of errors. First, there was an increase in the proportion of subjects who failed the first session, but passed the second session. This increase was a result of 4 subjects who made one or two errors on the first trial where they confused yellow and green with each other, but they either did not make any errors or it was an orange-red mistake at the second session. Second, there was an increase in the number of subjects who passed the first, but failed the second session. This was due to one subject who had no errors on the first session, but made one green-yellow error on the second session.

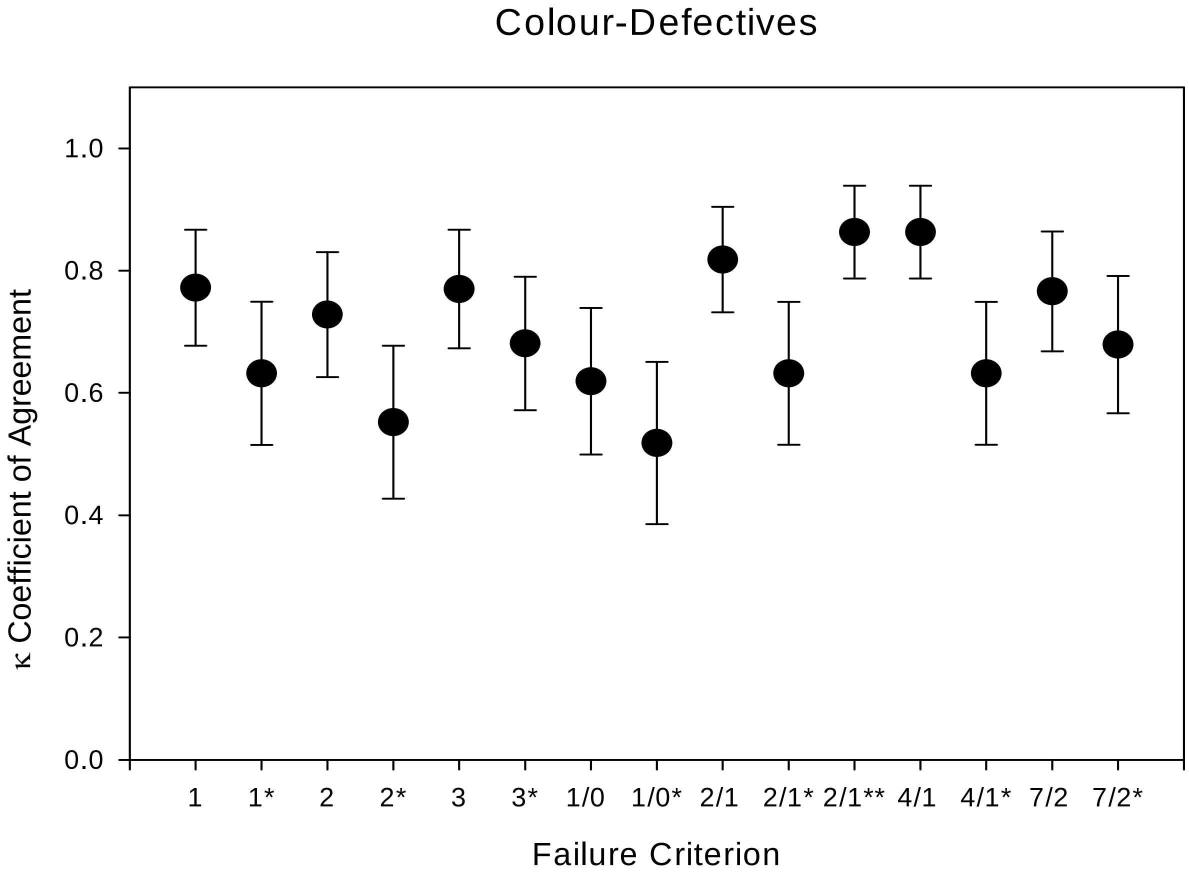


Figure 5.7. Kappa (κ) coefficients of agreement for repeatability for the colour-defective group using the failure criteria listed on the x-axis. The criteria defined in Table 5.4. Error bars represent ±1 standard error.

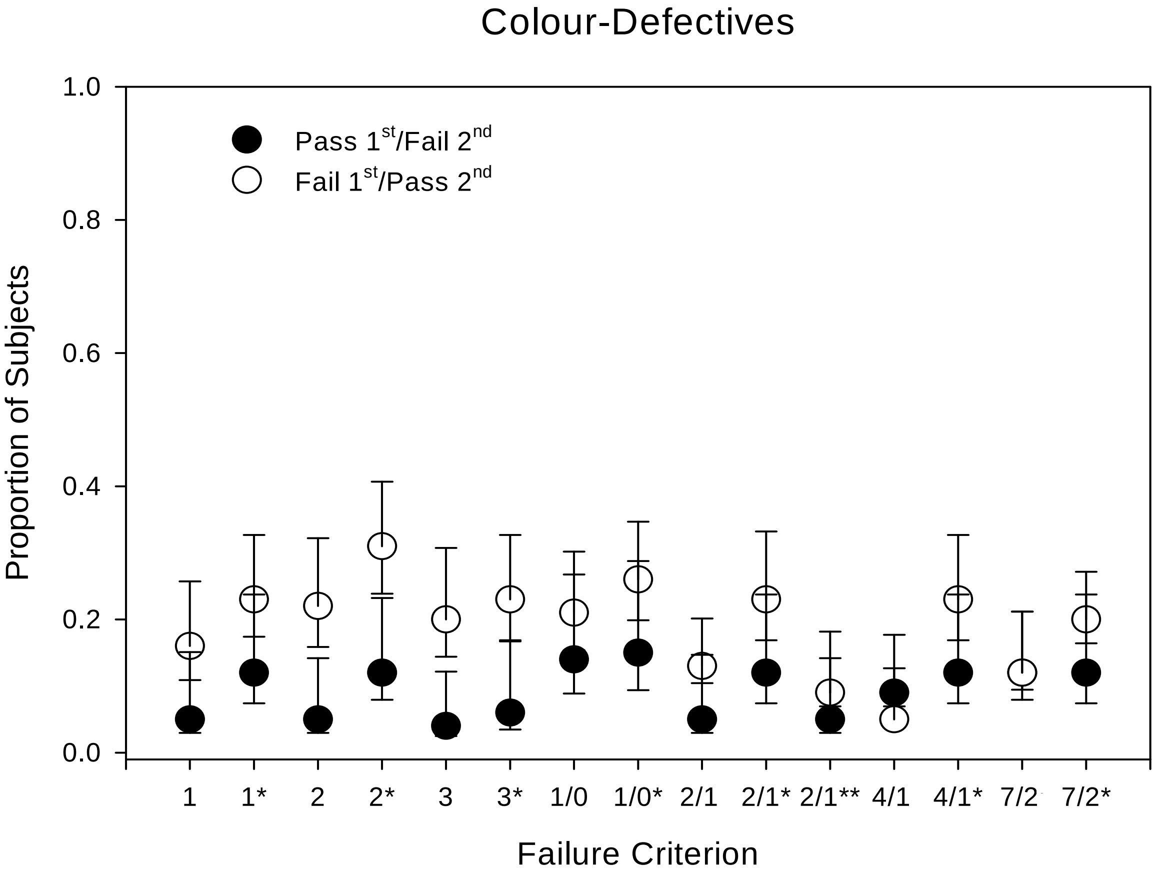


Figure 5.8. Frequency of the between-session discrepancies for the colour-defective group using the different failure criteria listed on the x-axis. The criteria defined in Table 5.4. Error bars represent ±1 standard error. These values were calculated based on a binomial distribution.

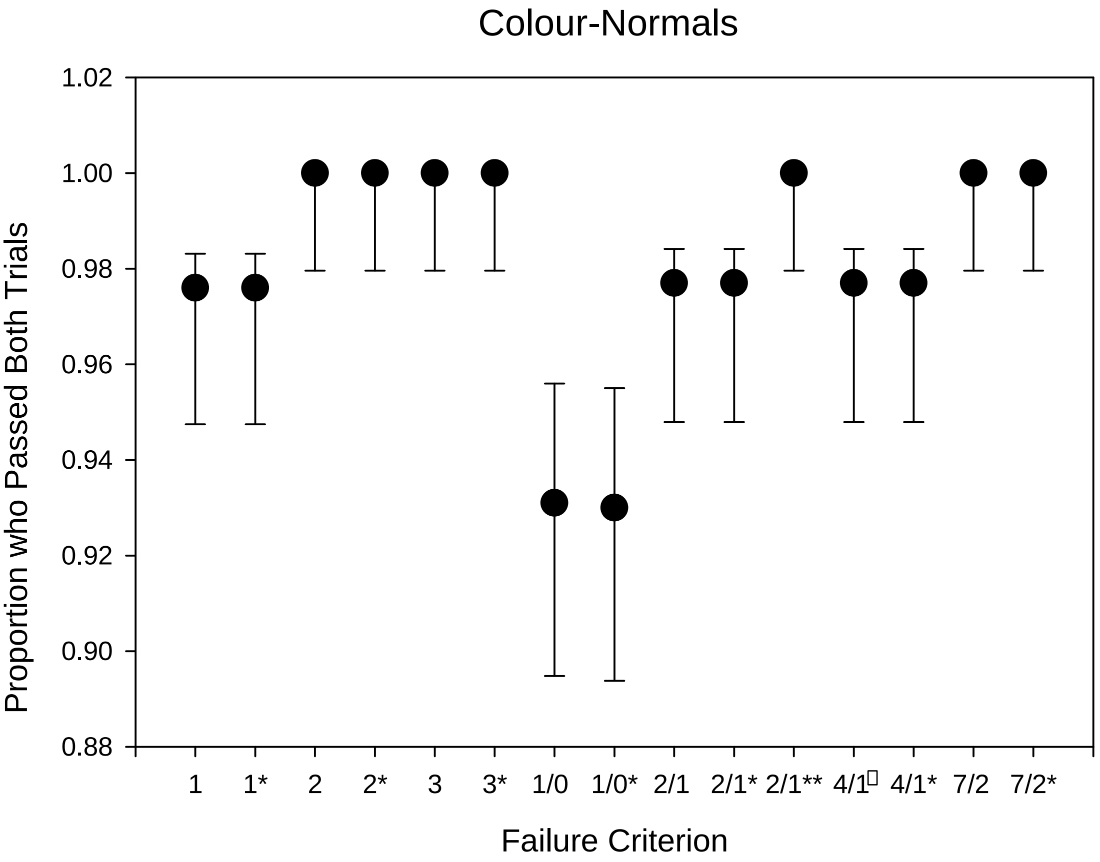


Figure 5.9. Frequency of the colour-normals who passed both sessions using the different failure criteria listed on the x-axis. The criteria defined in Table 5.4. Error bars represent $\underline{+}1$ standard error. These values were calculated based on a binomial distribution.

5.3.2. Scoring Matrix for the Various Criteria.

Given that a number of different parameter examined for establishing the pass/fail criteria, a scoring matrix (shown in Table 5.5) was set up. A value of 1.0 was assigned to the matrix cell if the scoring criterion met the condition listed in the column heading; if not, then zero was assigned. The total scores are given in the last column. There are several criteria which had values that were statistically identical to the best values based on the 95% confidence intervals. In order to establish statistical identity, however, we used a stricter measure. All values that were within one standard error of the best value were considered to be statistically identical to this value and so a value of 1.0 was entered into the matrix in these cases. The reason for using the standard error instead of the confidence intervals was that the stricter definition of statistically identical would be a more efficient method of finding the optimum pass/fail criterion based on the repeatability of the test.

The Table 5.5 shows that there was only one pass/fail criterion which met all the four conditions. This was to ignore any mistakes where orange was misnamed as red and then allow 2 errors on the first session and 1 error on the second session (the 2/1** criterion). However, allowing 4 errors on the first trial and one error on the second trial (i.e., 4/1) was very close. The disadvantage of the 4/1 criterion was that the colour-normal repeatability for passing both sessions was marginally worse than 2/1** criterion. Because the repeatability of two criteria was nearly identical for the colour-defective group, they will be examined more closely.

Table 5.5. Scoring matrix for various failure criteria. The criteria are defined in Table 5.3. The failure criteria that meet the conditions listed in the other columns or have values that are within one standard error of the optimum value are assigned a value of 1.0.

Failure Criterion	κ coefficient of repeatability is maximized for colour-defectives	Proportion of colour-defectives who passed the 1st trial, but failed the 2nd trial is minimized	Proportion of colour-defectives who failed the 1st trial, but failed the 2nd trial is minimized	Proportion of colour-normals who passed both sessions is maximized	Total Score
1	0	1.0	1.0	0	2.0
1*	0	1.0	0	0	2.0
2	0	1.0	0	1.0	2.0
2*	0	1.0	0	1.0	2
3	0	1.0	1.0	1.0	3.0
3*	0	1.0	0	1.0	2.0
1/0	0	0	1.0	0	1.0
1/0*	0	0	0	0	0
2/1	1.0	1.0	1.0	0	3.0
2/1*	0	1.0	0	0	1.0
2/1**	**1.0**	**1.0**	**1.0**	**1.0**	**4.0**
4/1	1.0	1.0	1.0	0	3.0
4/1*	0	1.0	0	0	1.0
7/2	0	1.0	1.0	1.0	3.0
7/2*	0	1.0	0	1.0	2.0

2/1**		Colour-Defectives First Trial	
		Pass	Fail
Second Trial	Pass	21	2
	Fail	1	20
κ =0.863 se= 0.075		Proportion Pass 1st but Fail 2nd =0.045	Proportion Fail 1st but Pass 2nd =0.091

4/1		Colour-Defectives First Trial	
		Pass	Fail
Second Trial	Pass	21	1
	Fail	2	20
κ =0.863 se= 0.075		Proportion Pass 1st but Fail 2nd =0.087	Proportion Fail 1st but Pass 2nd =0.048

Figure 5.10. Between-session pass/fail contingency tables for the colour-defective group for two different scoring methods. The left table is for a criterion where all orange misnamed as red errors are ignored and then 2 two errors are allowed on the first trial and 1 error is allowed in the second trial. The scoring criterion for the right table is 4 errors are allowed on the first trial and 1 error is allowed on the second regardless of the type of error.

Figure 5.10 shows the contingency tables for the colour-defective subjects using these two pass/fail scores. (The set of results for the entire list of failure criteria in Table 5.5 are in Appendix 5.1). The figure shows that the difference between the two criteria was a small difference in their between-session discrepancies. There was a slightly smaller probability of passing the first session and then failing the second session using the 2/1** criterion compared to the 4/1 criterion. This was also found in the colour-normal result.

It is unclear which of these two criteria would be more advantageous to use because two of the three discrepancies in each case are not the same subjects. The one common subject was an individual who passed the first session with zero errors and made three errors at the second trial. Two of these errors were typical of colour-normals (dark green misnamed as green) and one was not typical (misnaming red as orange). The other two discrepancies using the 2/1** criterion had three errors on the first trial and one on the second trial. On the first trial, both subjects had two of the errors that were typical of colour-normals and one that was typical of a colour-defective. At the second trial, they either had one colour-normal error or no errors. These two individuals passed both sessions using the 4/1 scoring method.

One of the discrepant results using the 4/1 method was a subject who had three errors on the first trial and did worse on the second session with 7 errors. The types of errors on both sessions were a mix of colour-normal and colour-defective mistakes. He failed both sessions using the 2/1** criterion. The second discrepant result using the 4/1 scoring method was a person who had 6 errors (3- orange was misnamed as red errors) on the first session and no errors on the second session.

Although the difference between the two criterion was small, is better from a safety perspective to have a higher percentage of individuals who failed at the first session and passed at the second session using the 2/1** than to have a higher percentage of individuals who passed the first session, but failed the second session which occurred with the 4/1 criterion. The one subject who went from 3 errors on the first session to 7 errors on the second also suggests that the 2/1** would be better to use from a safety perspective.

However, the question that is not resolved is whether failure on the VDT Dispatch test should also be made conditional on the types of errors for those colour-defectives who make a small number of errors that are atypical of colour-normals. Because of this possibility, we examined the responses of the all the 22 subjects who passed the first session, in particular the 9 who subjects made one or two mistakes at the first session. Of these 9 individuals (who passed), 3 (33%) had at least one error that was typical for colour-defective subject (confusing green and yellow with each other). None of these 3 individuals had any errors on the second session. The other 6 subjects however made 1 or 2 errors (atypical of colour defectives) on the first session and 1 (atypical of colour defectives) or no errors in the second session. Of the other 13 subjects who passed the first session, 2 (15%) had zero errors on the first session and made at least one colour-defective error on the second session. The remaining 11 subjects had zero errors in both the sessions with no colour-defective errors. Comparing the results from two sessions shows that the small number of colour-defective errors made in the first session were usually not repeatable, suggesting that with practice the probability of committing the errors decreases.

A general interpretation of this type of performance when a small number of errors is made is that one should expect a small number of non-repeatable errors on colours that are perceptually similar, such as yellow and green for the colour-defectives or orange and red for the colour-normals. Experience and training can reduce these types of errors if the error is pointed out early in the training. Nevertheless, the fact that these errors also occurred for the first time for some subjects on the second trial suggests that even if the person has a perfect score, they should be counselled that they are prone to making certain types of errors.

5.4. Conclusion

The VDT test was developed to determine whether a colour-defective person can identify the colours used to code information in the centralised traffic control display. The optimum pass/fail criterion was derived based on the repeatability of the total number and types of errors made by colour-normals. The difficulty in selecting this criterion was that a small percentage of colour-normals, made several errors. Some mistakes were probably due to data entry errors (example blue-green called blue), whereas the orange misnamed as red was due to the fact that these two colours were similar in appearance. Often, these colour-normal errors were not repeatable. The second difficulty was that some colour-defectives also had a small number of errors, but the types of the errors were indicative of a colour vision defect and atypical of the colour-normal errors. These errors were also not usually repeatable. Third, there was a significant learning effect for both groups. This learning effect was probably due to the fact that the subjects received feedback at the end of the first session as to the types of errors made on the test.

The first issue of the colour-normal errors was overcome by disregarding the orange misnamed as red (however, the red misnamed as orange is an error) and allowing at least 2 other errors on the test for the first session. The second problem of the differences in the types of errors made by colour-defectives is not considered in this scoring criterion. The issue of the learning component was taken into account by reducing the number of errors allowed of the second trial to 1 error (again, the orange-misnamed as red is discounted). Based on this criteria (2/1**) 58% of the colour-defectives (including all dichromats) failed the VDT test.

The next chapter will examine the types of errors on the VDT test in more detail.

References

1. Hays, W. (1981). “Analysing Qualitative Data: Chi-Square Tests. Chapter 15. Statistics”. New York, CBS College Publishing**:** 558-559.
2. Feinstein, A. R. and D. V. Cicchetti (1990). "High agreement but low kappa: I. The problems of two paradoxes." J Clin Epidemiol 43(6): 543-9.

CHAPTER 6

Results II. Types of Errors made by Colour-Defectives in the VDT Test

6.1. VDT Test Colours and Errors

The previous chapter showed that colour-defective observers made numerous mistakes on the Track Grid (second part) and Occupancy (third part) of the VDT test. This chapter will examine the types of errors in more detail by segregating the colour-defective group into the type of red-green defect and whether they pass or failed the VDT test. Only the first trial is examined because of the larger number of subjects.*

6.2. Types of Errors made by Colour-Defectives who Passed the VDT Test

Tables 6.1 through 6.10 present the types of errors made by the different color-defective groupings. The errors are given in the percentages relative to the total number of errors made by all subjects within a specific group for a given colour combination. The actual numbers of errors on each combination are listed in Appendix 6.1. The colour names on the left column are the actual colours and the colour names on the top are the subject's response. The shaded cells denote which colour confusions were predicted based on the dichromatic lines of confusion, ignoring luminance differences.

---

* A cursory look at the second session indicated that the same general trends in the types of errors were present.

Table 6.1 presents the types of errors made in the Track Grid (second part) and Occupancy (third part) of the VDT test by the colour-defectives (n=24) who passed the test. Two errors (excluding orange called red errors) in total were allowed for a pass.

The results show that the colour-defectives who passed the test generally made errors that were similar to colour-normals (Chapter 5.Table 5.1a). However, there was a slightly higher presentage of green-yellow errors in third part which would be expected based on their colour vision deficiency. Tables 6.2 and 6.3 show the results with this group divided into protanomalous and deuteranomalous subjects who passed the VDT test. The results show that the protans are less likely to confuse orange with red but more likely to confuse green and yellow than deutans.

6.3. Types of Errors made by Colour-Defectives who Failed the VDT Test

Table 6.4 presents the types of errors made in the Track Grid (second part) and Occupancy (third part) of the VDT test by the colour-defectives (n=28) who failed the test. This group includes all the nine dichromats. The general finding was that the number of errors was higher, but the percentage of errors on colours that were common to both parts was similar. As expected, the more frequent errors occurred for colours that were on, or near, the lines of confusion.

Table 6.5 to 6.10 show the types of errors made (in percentage) in the Track Grid (second part) and Occupancy (third part) of the VDT test for the colour-defectives who failed cateragorized by protans (combined anomalous trichromats and dichromats), deutans, protanopes, protanamalous, deuteranopes and deuteranamalous subjects. Note that comparisons of absolute numbers between the different categories should not be made because the number of subjects in the different categories was unequal. Nevertheless, reviewing the tables shows that percentages show two general trends.

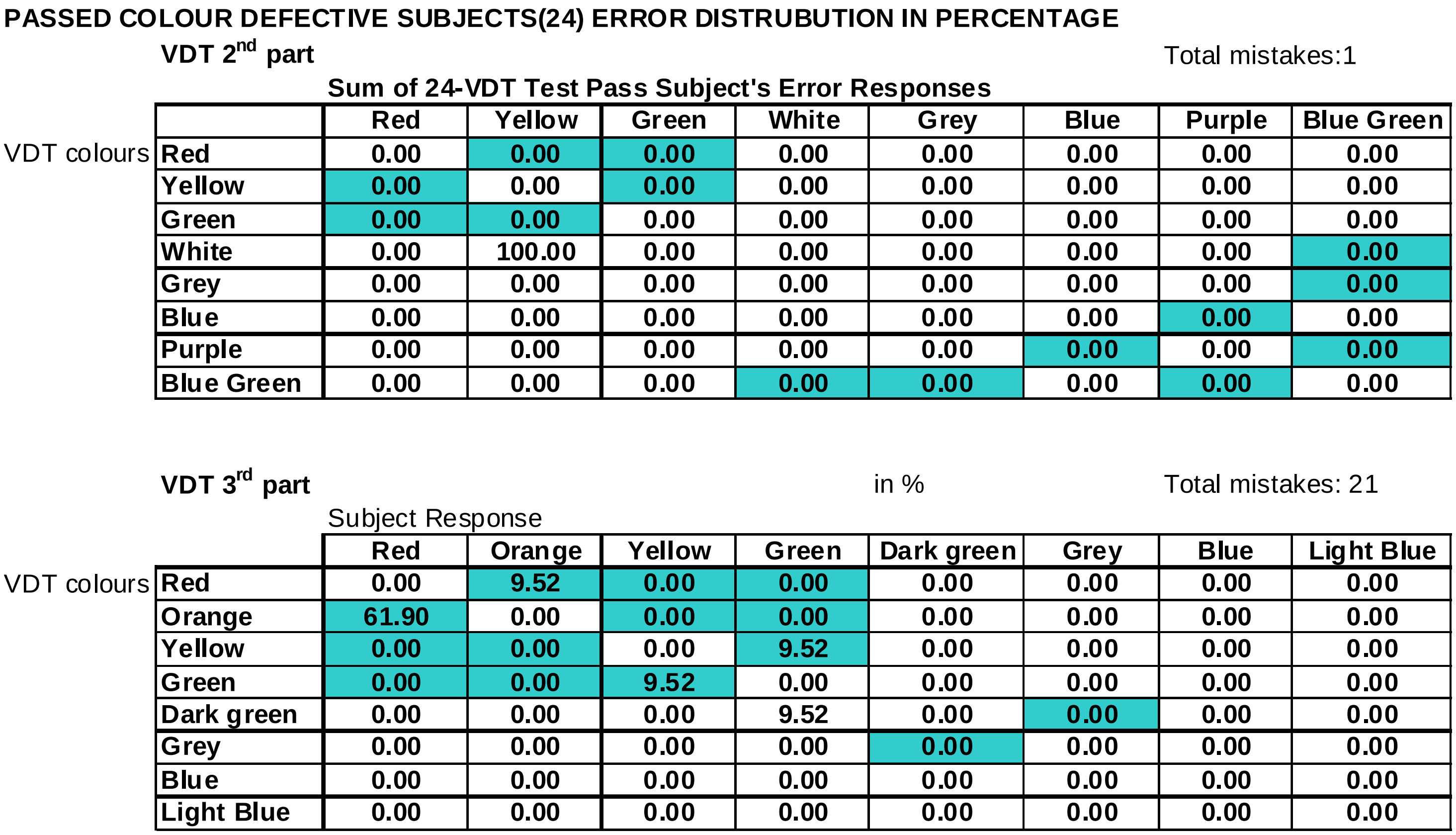
The first is that percentage of errors for the protans on each part was approximately equal. This result was due to the increase in the percentage of errors made by protanomalous subjects balanced by the decrease in errors made by protanopic subjects. The deuteranomalous subjects also showed an increase in the number of errors on the third part, but the increase was much larger than found for the protanomalous subjects. In contrast, the deuteranopes had only a slight decrease in errors on the third part so that the net result was a 55% increase in errors for the deutan group on the third part.

The other trend that emerged was the asymmetry in the nature of the errors. If two colours were likely on the same line of confusion, then one would expect the percentage of errors for each to be approximately equal. For example, the percentage of green misnamed as yellow errors should be approximately equal to the percentage of yellow misnamed as a green error. However, that was not the case. In the failed colour-defective group, the frequency that green was called yellow was 3 times more frequent than yellow called green for both the second and third parts. This asymmetry was present in all groupings with the deuteranomalous subjects showing the largest asymmetry and protanomalous subjects exhibiting the smallest asymmetry in these two errors. Another notable asymmetrical error combination was blue-purple errors in the second part. Purple called blue errors were 3 times more than the blue called purple errors in the second part. The deuteranomalous subjects, again, had the largest asymmetry.

The third asymmetry was the orange-red errors. Orange was called red 12 times (table 6.4) more than the red called orange errors in the third part. Again the asymmetry was must pronounced for the deuteranomalous subjects and least pronounced for the protanomalous individuals. This asymmetry in the red-orange errors was also present in the colour-normal and colour-defectives who passed results.

In general terms, the errors made by colour-defective subjects were consistent with the location of the colours with respect to the dichromatic lines of confusion. However, these predictions ignored luminance differences between the colours which could explain why some coloured pairs have a lower percentage of errors than others and the asymmetry present in errors. The next section examines the errors in more detail to determine whether there is a more quantitative method for determining when the errors will occur and whether the asymmetries in the nature of the errors can be explained.

Table 6.1. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by twenty four colour-defective subjects those who passed the VDT test (first trial). Note here that there were no dichromatic results.



PASSED COLOUR DEFECTIVE SUBJECTS(24) ERROR DISTRUBUTION IN PERCENTAGE

VDT 2nd part Total mistakes:1

Sum of 24-VDT Test Pass Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Purple	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

VDT 3rd part in % Total mistakes: 21

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	9.52	0.00	0.00	0.00	0.00	0.00	0.00
Orange	61.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	9.52	0.00	0.00	0.00	0.00
Green	0.00	0.00	9.52	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	9.52	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.2. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by six protanomalous subjects those who passed the VDT test (first trial). Note here that there were no dichromatic results.

Protanomalous(=protan) errors only

VDT 2nd part

Total mistakes: 0

Sum of 6 Passed Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Purple	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

VDT 3rd part

Total mistakes: 3

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Orange	33.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	33.33	0.00	0.00	0.00	0.00
Green	0.00	0.00	33.33	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.3. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by eighteen deuteranomalous subjects those who passed the VDT test (first trial).Note here that there were no dichromatic results.

Deuteranamalous(=deutan) errors only 19

VDT 2nd part Total mistakes: 1

Sum of 18 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Purple	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue Green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

VDT 3rd part Total mistakes: 18

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	9.52	0.00	0.00	0.00	0.00	0.00	0.00
Orange	57.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	4.76	0.00	0.00	0.00	0.00
Green	0.00	0.00	4.76	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	9.52	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.4. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by twenty eight colour-defective subjects those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS(28) ERROR DISTRUBUTION IN PERCENTAGE

VDT 2nd part

Total mistakes:178

Sum of 28 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	10.11	0.00	0.00	0.00	0.00	0.00
Green	0.00	32.02	0.00	0.00	0.00	0.00	0.56	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	1.12	0.00	0.56	0.56	0.00	0.00	4.49	8.99
Blue	0.00	0.00	0.00	0.00	0.00	0.00	3.93	0.00
Purple	0.00	0.00	0.00	0.00	0.00	11.80	0.00	2.81
Blue Green	0.00	0.00	2.81	1.12	12.36	2.25	4.49	0.00

VDT 3rd part

Total mistakes: 203

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	1.97	0.00	0.00	0.00	0.00	0.00	0.00
Orange	25.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.49	0.00	11.33	0.00	0.00	0.00	0.00
Green	0.00	2.96	31.53	0.00	0.49	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	8.37	0.00	8.87	0.00	0.00
Grey	0.00	0.00	0.00	0.49	7.88	0.00	0.00	0.49
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.5. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by fifteen protans those who failed the VDT test (first trial).

Total Protan errors

VDT 2nd part

Total mistakes: 101

Sum of 15 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	13.86	0.00	0.00	0.00	0.00	0.00
Green	0.00	32.67	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.99	0.00	0.99	0.99	0.00	0.00	1.98	15.84
Blue	0.00	0.00	0.00	0.00	0.00	0.00	4.95	0.00
Purple	0.00	0.00	0.00	0.00	0.00	10.89	0.00	0.00
Blue Green	0.00	0.00	0.00	1.98	14.85	0.00	0.00	0.00

VDT 3rd part

Total mistakes: 102

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.00
Orange	15.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	20.59	0.00	0.00	0.00	0.00
Green	0.00	1.96	39.22	0.00	0.98	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	6.86	0.00	4.90	0.00	0.00
Grey	0.00	0.00	0.00	0.00	8.82	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.6. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by thirteen deutans those who failed the VDT test (first trial).

Total Deutan errors

VDT 2nd part

Total mistakes: 77

Sum of 13 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	5.19	0.00	0.00	0.00	0.00	0.00
Green	0.00	31.17	0.00	0.00	0.00	0.00	1.30	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	1.30	0.00	0.00	0.00	0.00	0.00	7.79	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	2.60	0.00
Purple	0.00	0.00	0.00	0.00	0.00	12.99	0.00	6.49
Blue Green	0.00	0.00	6.49	0.00	9.09	5.19	10.39	0.00

VDT 3rd part

Total mistakes: 119

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	2.52	0.00	0.00	0.00	0.00	0.00	0.00
Orange	29.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.84	0.00	1.68	0.00	0.00	0.00	0.00
Green	0.00	3.36	20.17	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	8.40	0.00	10.92	0.00	0.00
Grey	0.00	0.00	0.00	0.84	5.88	0.00	0.00	0.84
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.7. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by five protanopes those who failed the VDT test (first trial).

Protanope errors only

VDT 2nd part Total mistakes:39

Sum of 5 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	5.13	0.00	0.00	0.00	0.00	0.00
Green	0.00	23.08	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	2.56	25.64
Blue	0.00	0.00	0.00	0.00	0.00	0.00	7.69	0.00
Purple	0.00	0.00	0.00	0.00	0.00	15.38	0.00	0.00
Blue Green	0.00	0.00	0.00	0.00	20.51	0.00	0.00	0.00

VDT 3rd part Total mistakes: 27

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Orange	29.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	18.52	0.00	0.00	0.00	0.00
Green	0.00	7.41	37.04	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	7.41	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.8. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by ten protanomalous those who failed the VDT test (first trial).

Protanomalous errors only

VDT 2nd part

Total mistakes: 62

Sum of 10 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	19.35	0.00	0.00	0.00	0.00	0.00
Green	0.00	38.71	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	1.61	0.00	1.61	1.61	0.00	0.00	1.61	9.68
Blue	0.00	0.00	0.00	0.00	0.00	0.00	3.23	0.00
Purple	0.00	0.00	0.00	0.00	0.00	8.06	0.00	0.00
Blue Green	0.00	0.00	0.00	3.23	11.29	0.00	0.00	0.00

VDT 3rd part

Total mistakes: 75

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	1.33	0.00	0.00	0.00	0.00	0.00	0.00
Orange	10.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	21.33	0.00	0.00	0.00	0.00
Green	0.00	0.00	40.00	0.00	1.33	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	9.33	0.00	6.67	0.00	0.00
Grey	0.00	0.00	0.00	0.00	9.33	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.9. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by four deuteronopes those who failed the VDT test (first trial).

Deuteranope errors only

VDT 2nd part

Total mistakes: 60

Sum of 4 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	6.67	0.00	0.00	0.00	0.00	0.00
Green	0.00	23.33	0.00	0.00	0.00	0.00	1.67	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	1.67	0.00	0.00	0.00	0.00	0.00	10.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	3.33	0.00
Purple	0.00	0.00	0.00	0.00	0.00	16.67	0.00	6.67
Blue Green	0.00	0.00	0.00	0.00	10.00	6.67	13.33	0.00

VDT 3rd part

Total mistakes: 45

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Orange	24.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Green	0.00	8.89	22.22	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	0.00	0.00	24.44	0.00	0.00
Grey	0.00	0.00	0.00	2.22	15.56	0.00	0.00	2.22
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.10. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test (expressed in percentage) by nine deuteranomalous subjects who failed the VDT test (first trial).

Deuteranamalous errors only

VDT 2nd part

Total mistakes:17

Sum of 9 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Green	0.00	58.82	0.00	0.00	0.00	0.00	0.00	0.00
White	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Purple	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.88
Blue Green	0.00	0.00	29.41	0.00	5.88	0.00	0.00	0.00

VDT 3rd part

Total mistakes: 56

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	5.36	0.00	0.00	0.00	0.00	0.00	0.00
Orange	42.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Yellow	0.00	1.79	0.00	3.57	0.00	0.00	0.00	0.00
Green	0.00	0.00	25.00	0.00	0.00	0.00	0.00	0.00
Dark green	0.00	0.00	0.00	17.86	0.00	3.57	0.00	0.00
Grey	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Light Blue	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

6.4. A Closer Look at the Types of Errors and Colour Confusions

The colour confusions were generally consistent with previous research and characteristics of the colour vision defect. The issue examined in the section is how well the chromaticity lines of confusion will predict the errors when there are luminance clues present. Figures 6.1 and 6.2 show the VDT test colours with respect to the protanopic and deuteranopic lines of confusion. The expected colour confusions for protans and deutans are listed along with the averaged error rates in Table 6.11, 6.12, 6.13 and 6.14. For the present analysis, the luminance difference has been ignored, but they are listed for comparison. The luminance ratios were calculated based on the luminance measured for a given colour with the Minolta CS-100 photometer. The relative luminance ratios for protanopes and deuteranopes in respective dichromatic colour space were calculated from the sum of the transformed L and M values described in section 6.5.2.

The asymmetries in the types of errors for a given pair have been ignored and the error rates for each confusion pair in the tables are the average rate for the two possible types of mistakes.

The distribution of errors on the different colour combinations were presented in the previous chapter. In order to compare the colour confusions with errors made per colour combination per person, we need to find the errors made by each group for a given colour combination and the number of times the colour was presented. Hence the error rates in the following section are calculated differently from the previous section. For example, the number of green called yellow errors for protans who failed the VDT test was 33 and yellow called green errors was 14 (Table. G. Appendix 6.1). The green called yellow error frequency was calculated based on the total number of errors (33) made for that colour combination divided by the product of the total number of colour presentation (6) and the number of subjects (15). This value (0.37) was expressed in percentage. Similarly, the yellow called green error frequency was calculated based on the total number of errors (14) made in that colour combination box divided by the total number of colour presentation (6) multiplied by the number of subjects (15). This value (0.16) was converted into percentage. Finally, the green called yellow error percentage (37) and yellow called green percentage (16) was averaged. This value (≃28%) gives the frequency of error percent per person for a given colour combination for protans in Table 6.11.

6.4.1. Second Part Colour Confusions and Errors

Protan confusions: There are seven pairs of potential colour confusions based on figure 6.1 that could occur for the protans in the second part of the VDT test. Table 6.11 shows that the protanomalous subjects who passed did not make any errors consistent with the lines of confusion. For the protans who failed the yellow-green errors were the most frequent errors for the protans (18%) followed by blue green-grey, purple-blue, white-blue green and white-grey. Assuming, the subjects were guessing, then one would expect that the error rates would vary from 25% to 50% depending on the number of colours falling in the lines of confusion. For example, because green could be confused with yellow and red in the second part of the test, the percentage of green called yellow errors would be close to 33%. In the third part the percentage of green called yellow errors would be approximately 25% because orange is also on the same line of confusion. Of course, this assumes that the subject is a dichromat and luminances are equal. Notice that the percentage of errors rarely approached the guess rate (not even for the dichromats). If it did, for example the 28% error rate for the green-yellow errors, there was not the corresponding high error rate for the green-red and yellow-red errors. In fact there were no errors made on green-red and yellow-red combinations. This was probably due to the high luminance difference between these colours that aided in correct identification.

Deutan confusions: There were eight pairs of potential colour confusions (based on figure 6.1) that could occur for the deutans in the second part of the VDT test. Seven of the eight pairs were similar as protan confusion colours. The additional confusion pair was the blue green-purple combination which had about a 3% error rate. Similar to protans, the deuteranomalous subjects who passed did not make any errors consistent with the lines of confusion. For the deutans who failed, their results were also similar to protan results with yellow-green errors as the most frequent error. Blue-green- purple errors and blue-purple errors were the next most common for the deutans. Unlike the protanopes, the deuteranopes always had a higher error rate than their anomalous trichromatic counterparts (Although not shown in Table 6.12, this could be noted in Table 6.9 & 6.10).

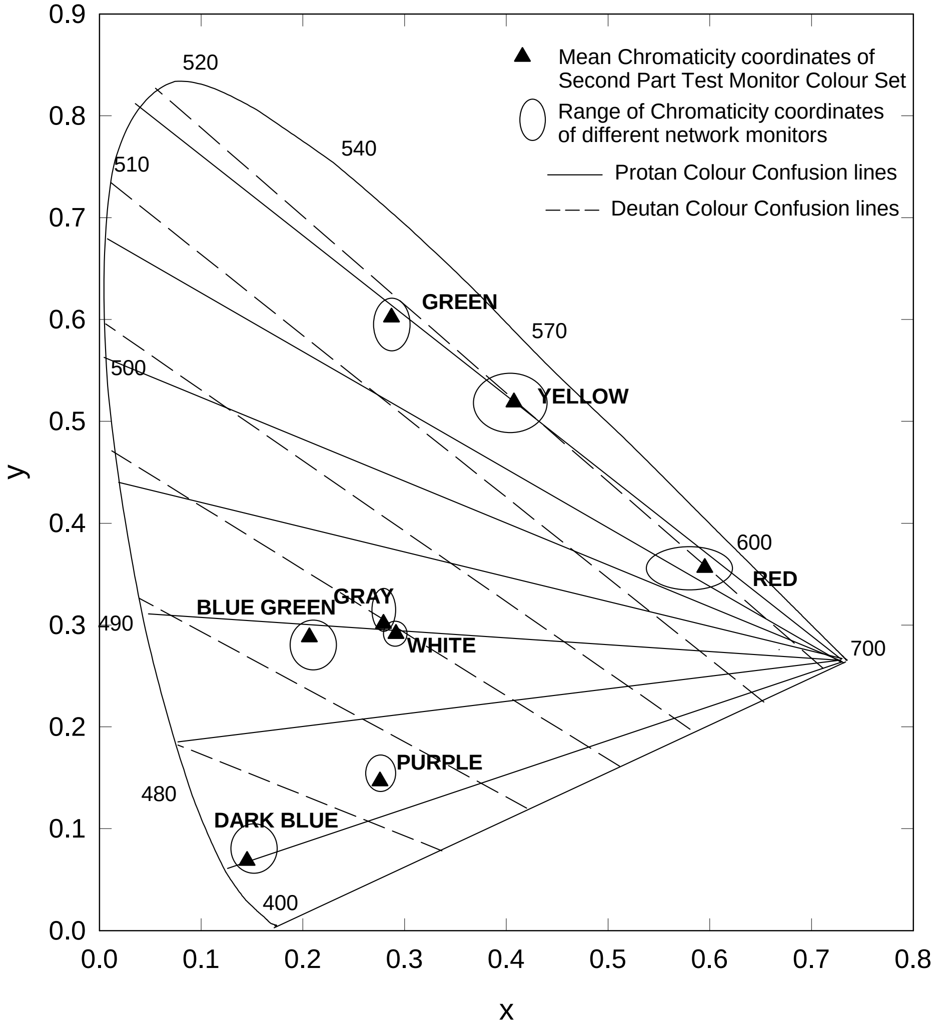


Figure 6.1. Mean chromaticity coordinates of the second part test colours (solid filled triangles) in the 1931 CIE chromaticity diagram (Appendix 6.2). The ellipses represent the range of chromaticity coordinates measured for various monitors at the Network Management Centre. The solid and dashed radiating lines represent the protanopic and deuteranopic lines of confusions.

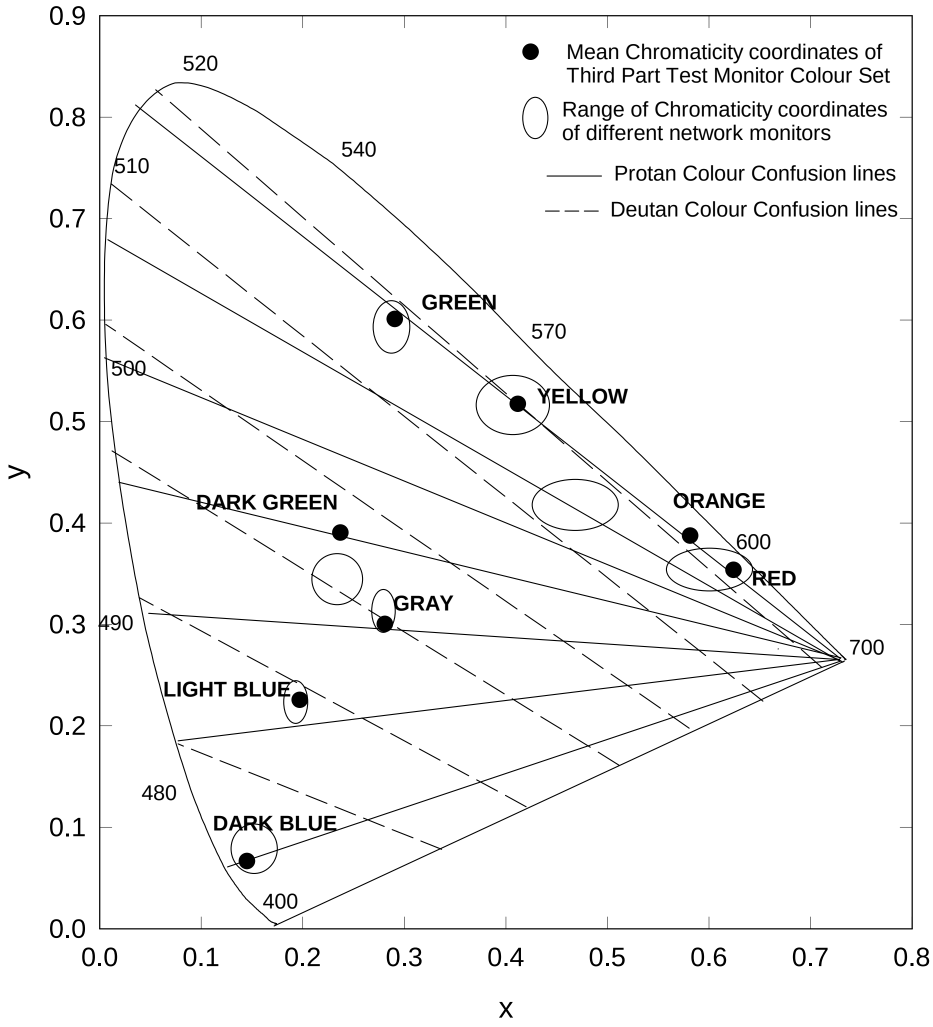


Figure 6.2. Mean chromaticity coordinates of the third part test colours (solid filled circles) in the 1931 CIE chromaticity diagram (Appendix 6.2). The ellipses represent the range of chromaticity coordinates measured for various monitors at the Network Management Centre. The solid and dashed radiating lines represent the protanopic and deuteranopic lines of confusions.

Table 6.11. Potential protan confusion colours and the frequency of errors for the second part of the VDT test.

Potential colour confusion **Second part** Only one combination for a pair is listed below as the error rates were averaged for a given combination.	**Luminance Ratio** **(Measured)**	**Relative Luminance Ratio in Protanopic colour space**	**Frequency of errors (in%) for a particular colour combination.** **(Averaged across the two possible types of errors)**			
Colour-Subject's Response	**Brighter Colour** Dimmer Colour	**Brighter Colour** Dimmer Colour	**Protans (21)**	**Protans who passed (6)**	**Protans who failed (15)**	**Protanopes (5)**
Yellow-Green	1.2	1.1	18%	0	28%	19%
Blue Green-Grey	3.6	3.8	12%	0	18%	30%
Purple-Blue	1.3	1.1	5%	0	9%	15%
White-Blue Green	1.9	1.7	1%	0	1%	0
Green-Red	3.3	6.6	0	0	0	0
Yellow-Red	4.1	7.0	0	0	0	0
White-Grey	6.5	6.6	0.5%	0	0	0.5%

Table 6.12. Potential deutan confusion colours and the frequency of errors for the second part of the VDT test.

Potential colour confusion **Second part** Only one combination for a pair is listed below as the error rates were averaged for a given combination.	**Luminance Ratio** **(Measured)**	**Luminance Ratio in Deuteranopic colour space**	**Frequency of errors (in%) for a particular colour combination** **(Averaged across the two possible types of errors)**			
Colour-Subject’s Response	**Brighter Colour** Dimmer Colour	**Brighter Colour** Dimmer Colour	**Deutans (31)**	**Deutans who passed (18)**	**Deutans who failed (13)**	**Deuteranopes (4)**
Yellow-Green	1.2	1.3	8%	0	18%	37%
Blue Green-Grey	3.6	3.4	2%	0	4%	12%
Purple-Blue	1.3	1.2	3%	0	8%	25%
White-Blue Green	1.9	1.8	0	0	0	0
Green-Red	3.3	2.5	0	0	0	0
Yellow-Red	4.1	3.3	0	0	0	0
Blue Green-Purple	3.6	2.7	3%	0	8%	25%
White-Grey	6.5	6.3	0	0	0	0

Table 6.13. Potential protan confusion colours and the frequency of errors for the third part of the VDT test.

Potential colour confusion **Third part** Only one combination for a pair is listed below as the error rates were averaged for a given combination.	**Luminance Ratio (Measured)**	**Relative Luminance Ratio in Protanopic colour space**	**Frequency of errors (in%) for a particular colour combination.** **(Averaged across the two possible types of errors)**			
Colour-Subject's Response	**Brighter Colour** Dimmer Colour	**Brighter Colour** Dimmer Colour	**Protans (21)**	**Protans who passed (6)**	**Protans who failed (15)**	**Protanopes (5)**
Yellow-Green	1.2	1.1	25%	3%	33%	25%
Orange-Red	3.1	3.8	7%	1.5%	9%	13%
Dark Green-Grey	2.5	2.9	5%	0	8%	3%
Orange-Green	1.1	1.8	1%	0	1%	3%
Yellow-Orange	1.3	1.9	0	0	0	0
Green-Red	3.3	6.6	0	0	0	0
Yellow-Red	4.1	7.0	0	0	0	0

Table 6.14. Potential deutan confusion colours and the frequency of errors for the third part of the VDT test.

Potential colour confusion **<u>Third part</u>** Only one combination for a pair is listed below as the error rates were averaged for a given combination.	**Luminance Ratio (Measured)**	**Relative Luminance Ratio in Deuteranopic colour space**	**Frequency of errors (in%) for a particular colour combination.** **(Averaged across the two possible types of errors)**			
Colour-Subject's Response	**<u>Brighter Colour</u>** Dimmer Colour	**<u>Brighter Colour</u>** Dimmer Colour	**Deutans (31)**	**Deutans who passed (18)**	**Deutans who failed (13)**	**Deuteranopes (4)**
Yellow-Green	1.2	1.3	7%	1%	17%	21%
Orange-Red	3.1	3.0	14%	6%	25%	23%
Dark Green-Grey	2.5	1.2	5%	0	14%	38%
Orange-Green	1.1	1.1	1%	0	3%	9%
Yellow-Orange	1.3	3.0	0.5%	0	0.5%	0
Green-Red	3.3	2.5	0	0	0	0
Yellow-Red	4.1	3.3	0	0	0	0

6.4.2. Third Part Colour Confusions and Errors

Protan confusions: There are seven pairs of potential colour confusions (based on figure 6.2) that could have occurred for the protans in the third part of the VDT test. Again, the yellow-green errors were the most frequent errors for the protans (25%) followed by orange-red, dark green-grey and orange-green. There were no errors made on green-red and yellow-red combinations presumably because the luminance differences aided the subject.

Deutan confusions: There are seven pairs of potential colour confusions (based on figure 6.2) that could have occurred for the deutans in the third part of the VDT test. They were similar as protan confusions. Although the trend was similar to the protan results, the orange-red errors were the most frequent errors for the deutans followed by yellow-green, dark green-grey and orange-green.

6.4.3. Discussion

The types of errors made by the colour-defectives who passed showed similar trend as the errors made by the colour-normals. The only exception to this trend was a small number of green called yellow mistakes in the Occupancy authority (third part) of the test. On the other hand, the types and number of errors made by the colour-defectives who failed were more pronounced and showed a trend that was consistent with their respective lines of confusion.

The number of errors made by the protans for the Track grid (second part) and Occupancy authority (third part) was approximately equal whereas the deutans had a 55% increase in errors in the Occupancy authority part. This was primarily due to the 2 fold increase in errors on the third part by the deuteranomalous subjects. The increase was primarily due to orange being identified as red. Interestingly the deuteranopes made this type of error, but it did not increase their total number of mistakes. They actually had a very slight reduction in the errors on the third part due to a reduction in yellow-green errors.

The other interesting result that emerged was the asymmetry on how the errors were made. If there was confusion between two colours, then one would expect the percentage of errors for each to be approximately equal. However, that was not the case. In the failed colour-defective group, the frequency that green was called yellow was 3 times more frequent than yellow called green for both the second and third parts. Other examples were, the purple was called blue 3 times more frequently than the blue was called purple in the second part and orange was called red 12 times more than the red was called orange errors in the third part. This asymmetry in the orange-red errors was also present in the results of colour-normal and colour-defectives who passed.

A possible reason for this asymmetry may be due to a response bias based on brightness contrast. For example, a colour-defective cannot decide whether a colour is green or yellow, but because the coloured object appears brighter relative to the surrounding colours, they will identify the colour as yellow because they have learned to associate yellow with being a brighter colour. Thus, there would be relatively few errors on the yellow stimulus itself unless it was surrounded by brighter objects. In this experiment, yellow was rarely surrounded by brighter stimuli since white was the only colour brighter than yellow. However, there should be relatively more errors on the green especially if the neighbouring objects were relatively dim. The dimmer objects would make the green appear relatively bright and so it would be identified as yellow.

This bias could also explain the asymmetry in blue green-grey errors, orange-red errors and purple-blue errors, where blue green was brighter than grey, orange was brighter than red and purple was brighter than blue. In these examples, the brighter colour of each pair was always dimmer than at least 40% of colours in the set. Thus, if these colours were surrounded by brighter colours they would appear dim and the person would mistake them for the dimmer colour of the pair.   Because of the randomization of the colours within and across the screens in each part, we did not have sufficient data to test this hypothesis quantitatively

Errors were more frequent between colours that fell near the same line of confusion and had a similar luminance. Nevertheless, the interaction between luminance and location with respect to the lines of confusion was not easy to interpret. One index as to how the luminance difference can reduce mistakes is the subject’s performance on the white-grey colours. These two achromatic colours have nearly the same chromaticity coordinates and differ only in luminance. The luminance ratio between the white and grey is approximately 7. Thus one would expect no errors for colours on the same line of confusion if the luminance ratio is at least 7 between the two colours.

There were, however, other data suggesting that the luminance ratio could be smaller than 7 and the error rate would be near zero depending on the colours. In the second part, deuteranopes made no red-yellow errors which had a luminance ratio of 3; however, they did make many errors on the purple and blue-green which also had a luminance ratio of 3. The protanopes showed similar behaviour on other coloured pairs. They did not confuse white and blue-green even though the luminance ratio between these two colours was less than 2, but they did make numerous blue-green and grey confusions even though the luminance ratio between the two colours was larger.

One factor that could be influencing this type of behaviour is the average luminance of the coloured pairs. In the case of the deuteranopes with no red-yellow errors and multiple blue-green-purple errors, the average luminance of the red and yellow was 30.5 cd/m^2, whereas the average luminance of the blue-green and purple stimuli was 19 cd/m^2, Similarly for the protanopes, the average luminance of white and blue-green was 42 cd/m^2 and the average luminance of the grey and blue-green was 19 cd/m^2, It appears that a smaller luminance ratio is required for error-free performance when the average luminance is higher at least for this display.

This potential influence of luminance behaviour demonstrates that it is difficult to accurately predict performance based “eyeballing” the colours with respect to the lines of confusion and taking into account luminance differences. Nevertheless, examining where the colours fall with respect to the lines of confusion and the luminance differences does provide a very crude rule that if the colours near the same line of confusion and the colour normal luminance ratio is less than 7, then errors could occur. This is especially true if the colours are relatively close together in the chromaticity diagram and relatively dim. If the colour normal luminance ratio is greater than 7, then errors for colours near the same lines of confusion are unlikely. For some colours a luminance ratio of 3 to 4 may be sufficient, especially if they are farther apart in the diagram and relatively bright.

CHAPTER 7

Results III. VDT Test Colour Differences and Error Correlation

7.1. VDT Test Colour Differences and Error Correlation

The previous chapter showed the error rates for the colours can be qualitatively predicted based on the location of the colours with respect to the lines of confusion and luminance differences. This chapter will examine the relationship in a more quantitative manner to determine whether there is more accurate method for predicting when colour confusions are likely to occur.

The previous section showed that errors were more likely for colours near the same line of confusion that had a smaller colour difference in the CIE xyz diagram. This observation suggests that calculating colour differences (including luminance differences) between pairs of colours for a normal colour space may provide useful information, if not, provide a baseline for comparing colour differences in dichromatic colour spaces.

Colour differences (ΔE) between any two pairs of colours are usually determined by calculating the Euclidean distance between the pair for a given colour space. The spaces selected for this study are the 1976 CIE L*a*b* and LMS Normal and dichromatic cone spaces.

7.1.1. CIE L*a*b* Colour Space

The CIE L*a*b* space is a non-linear transformation of the XYZ space and is considered to be perceptually more uniform than the other linear transformations of the 1931 space when there are large differences in colour. It is the recommended [1] colour space for calculating colour differences that are generally several just-noticeable-differences (JNDs) apart. The L* component closely matches human perception of lightness (i.e. brightness), a* corresponds to the red-green dimension and b* corresponds to the yellow-blue dimension.

To calculate the colour difference in CIE L*a*b* space, the x, y, and luminance values (Y) measured for each colour using the Minolta CS-100 were converted to XYZ tristimulus values using the following equations.

- Y is the luminance
- $X = Yx/y$
- $Z=Yz/y$

The XYZ values of each colour were converted to L*a*b* coordinate using these equations

$L^*= 116(Y/Y_n)^{1/3}-16$

$a^*= 500[(X/X_n)^{1/3} -(Y/Y_n)^{1/3}]$

$b^*= 500[(Y/Y_n)^{1/3} -(Z/Z_n)^{1/3}]$

where X_n Y_n Z_n are the stimulus values for a reference white. In this study, the white colour presented in the experiment was used as the reference white.

The ΔE* is calculated as:

$$\Delta E^* = [(L_1^*-L_2^*)^2 + (a_1^*-a_2^*)^2 + (b_1^*-b_2^*)^2]^{0.5}$$

where subscript numbers represent two different colours. The ΔE* colour differences for different colour combinations are listed in Appendix 7.1 for reference

7.1.2. LMS Normal and Dichromatic Cone Space

The LMS space was based on the cone responses for both colour-normals and dichromats. The LMS for colour-normals is a three dimensional colour space represented by the response (sensitivity) of the three types of cones (long, medium and short wavelength sensitivity cones) in the retina. It's a linear transformation of the XYZ space. This would be a more physiological space relative to the CIE XYZ system. Dichromatic cone space for the respective dichromats were calculated using algorithm developed by Brettel et al, [Brettel et al 1997 #77] for transforming digitized colour images to simulate dichromatic vision for normal observers was used. This algorithm essentially reduces the volume of the cone space by collapsing colours to one long wavelength sensitive cone type and a short wavelength sensitive cone.

Colour Difference in normal LMS cone space:

To calculate the colour difference in Normal LMS cone space, the LMS values of the individual colours were calculated using the following equation for a 2 degree field.[2]

$L = 0.17156X + 0.52901Y - 0.02199Z$

$M = -0.15955X + 0.48553Y - 0.04298Z$

$S = 0.01916X - 0.03989Y + 1.03993Z$

The ΔE_N was calculated as:

$$\mathbf{\Delta E_N = [(L_1-L_2)^2 + (M_1-M_2)^2 + (S_1-S_2)^2]^{0.5}}$$

where subscript numbers represent two different colours.

Colour Difference in Dichromatic Cone Space

Colour stimuli were represented as vectors in three dimensional LMS space and the algorithm (expressed in the same space) replaced each stimulus onto a reduced stimulus surface. Neutral stimuli (e.g. white) for colour normals were assumed to be perceived as neutral for dichromats. Similarly, stimulus of 575nm (yellow) and 475nm (blue) were assumed to be perceived as same in protanopes and deuteranopes. These two wavelengths serve as anchors in order to establish the proper plane. The algorithm is as follows.

Q is stimulus in a colour normal LMS space. This colour can be projected into a plane defined by the stimuli E (neutral colour), the monochromatic anchor stimulus A, and the origin O . The point would be labelled Q’

For a given equal energy stimulus E (L_E, M_E, S_E) and anchor stimulus A (L_A, M_A, S_A) the linear equation for the coordinates $L_{Q'}$, $M_{Q'}$, and $S_{Q'}$ of stimulus Q’is:

$aL_{Q'}+bM_{Q'}+cS_{Q'}=0$

with,

$a=M_ES_A-S_EM_A$

$b=S_EL_A-L_ES_A$

$c=L_EM_A-M_EL_A$

where L_E, M_E and S_E are the vector component of equal energy white and

L_A, M_A and S_A are the vector component of the monochromatic anchor stimulus A

The equations to transform into **protanopic (^{P})** space for a given stimulus Q (L_Q, M_Q, S_Q) are

$^PL_{Q'} = -(bM_Q+cS_Q)/a$

$^PM_{Q'} = M_Q$

$^PS_{Q'} = S_Q$

If $S_Q/M_Q < S_E/M_E$, then $(\lambda)_A$ =575 nm; else $(\lambda)_A$= 475 nm

The luminance for protanopes = $^PL_{Q'} {}^{+}\, {}^PM_{Q'}$

The colour difference ΔE for protanopic cone space was calculated as:

$$\mathbf{\Delta E = [(^PL_{Q'1}-{}^PL_{Q'2})^2 + (^PM_{Q'1}-{}^PM_{Q'2})^2 + (^PS_{Q'1}-{}^PS_{Q'2})^2]^{0.5}}$$

where subscript numbers(1 & 2) represent two different colours.

The equations for **deuteranopic(^{D})** space for a given stimulus Q (L_Q, M_Q, S_Q) are

$^DL_{Q'} = L_Q$

$^DM_{Q'} = -(aL_Q+cS_Q)/b$

$^DS_{Q'} = S_Q$

If $S_Q/L_Q < S_E/L_E$, then $(\lambda)_A$ =575 nm; else $(\lambda)_A$= 475 nm

The luminance for deuteranopes = $^DL_{Q'} {}^{+}\, {}^DM_{Q'}$

The colour difference ΔE for deuteranopic cone space was calculated as:

$$\Delta E * = [({}^{D}L_{Q'1} - {}^{D}L_{Q'2})^2 + ({}^{D}M_{Q'1} - {}^{D}M_{Q'2})^2 + ({}^{D}S_{Q'1} - {}^{D}S_{Q'2})^2]^{0.5}$$

where subscript numbers (1 & 2) represent two different colours.

The ΔE colour difference that were calculated are for different colour combinations for normal and dichromatic cone space are listed in Appendix 7.1.

7.1.3. Colour Differences and Average Errors Correlation

Table 7.1 & 7.2 lists the Pearson correlation coefficient along with the p value (for the Fisher exact statistical test) between the *average percent error* and *colour difference* ΔE* for the colour-normals, colour-defectives who passed, colour-defectives who failed and the dichromats in the CIE L*a*b* space, normal and dichromatic cone space. They are plotted in Figures 7.1a to 7.8a (Track Status) and Figures 7.1b to 7.8b (Occupancy Status).

Figures 7.1a to 7.8a show scatter plots of the *average percent error* as a function of *colour difference* ΔE* for the colour-normals, colour-defectives who passed, colour-defectives who failed and the dichromats in the CIE L*a*b* space. Figures 7.1b to 7.8b show scatter plots of the *average percent error* as a function of *colour difference* ΔE* for the colour-normals, colour-defectives who passed, colour-defectives who failed and the dichromats in the cone space. The solid lines are the linear correlation between the two variables and the curved lines are the nonlinear fits.

Table 7.1. Correlation between averaged error percent and colour difference in CIE L*a*b* space, Normal cone space and dichromatic cone space for the second part.

Second part	**CIE L* a* b* Space**		**Normal Cone Space**		**Dichromatic Cone Space**	
Groups	Linear Regression	Non Linear Regression	Linear Regression	Non Linear Regression	Linear Regression	Non Linear Regression
Colour-Normals	-0.03 (p=0.889)	NA	-0.24 (p=0.225)	NA	NA	NA
Passed- Colour Defectives	-0.07 (p=0.782)	NA	0.18 (p=0.349)	NA	0.12 (p=0.312)	NA
Failed- Colour Defectives	-0.37 (p=0.102)	0.36 (p=0.06)	-0.30 (p=0.117)	0.61 (p=0.0005)	-0.36 (p=0.104)	0.71 (p=0.0001)
Dichromats	-0.47 (p=0.034)	0.45 (p=0.01)	-0.31 (p=0.024)	0.47 (p=0.011)	-0.28 (p=0.017)	0.36 (p=0.0063)

Table 7.2. Correlation between averaged error percent and colour difference in CIE L*a*b* space, Normal cone space and dichromatic cone space for the third part.

Third part	**CIE L* a* b* Space**		**Normal Cone Space**		**Dichromatic Cone Space**	
Groups	Linear Regression	Non Linear Regression	Linear Regression	Non Linear Regression	Linear Regression	Non Linear Regression
Colour-Normals	-0.36 (p=0.058)	NA	-0.16 (p=0.406)	NA	NA	NA
Passed- Colour Defectives	-0.41 (p=0.032)	0.68 (p=0.0001)	-0.22 (p=0.254)	0.15 (p=0.043)	-0.22 (p=0.254)	0.29 (p=0.027)
Failed- Colour Defectives	-0.49 (p=0.007)	0.69 (p=0.0001)	-0.38 (p=0.045)	0.52 (p=0.004)	-0.38 (p=0.041)	0.63 (p=0.0001)
Dichromats	-0.49 (p=0.008)	0.70 (p=0.0001)	-0.40 (p=0.009)	0.56 (p=0.002)	-0.35 (p=0.004)	0.63 (p=0.0001)

The general observation from the linear correlations is that the errors decreased as the colour difference increased. However, the correlation was weak. The maximum correlation coefficient was -0.49 which indicates that the linear correlation could only predict 25% of the variance. The other trend to note was that correlations shown for the dichromatic transformations were no better than the CIE L*a*b* correlation which indicates that the transformations into dichromatic (or trichromatic) cone space did not offer any advantage over the ΔE*s calculated using the CIE L*a*b* formula even for the dichromatic data.

The result that the colour-normals and colour-defectives who passed did not have a significant correlation between the colour difference and error rate was not surprising because the number of errors was so small. On the other hand both these groups had a significant (or nearly significant) inverse correlation on the third part. This correlation between the ΔEs and error rate was primarily due to the orange-red errors.

The scatter plots also include a nonlinear (polynomial-inverse first order) curve fit to the data. This relationship provides a better description of the relationship between the ΔEs and error rate. Nevertheless, the fit is similar for both CIE L*a*b and cone colour difference formula which indicates that the dichromatic transformations provide little advantage over a more traditional trichromatic space. The result that the linear correlations between the ΔEs and error rates was weak was not surprising given that, although the perceptual steps are more uniform, they are still not equal. [1]

Figures.7.1.to 7.8.Correlation between averaged errors and colour difference in CIE Lab space (left side graphs) and normal and dichromatic cone space (right side graphs) for the second part(Track Status) of the VDT test for colour normals, passed and failed colour defectives and dichromats.

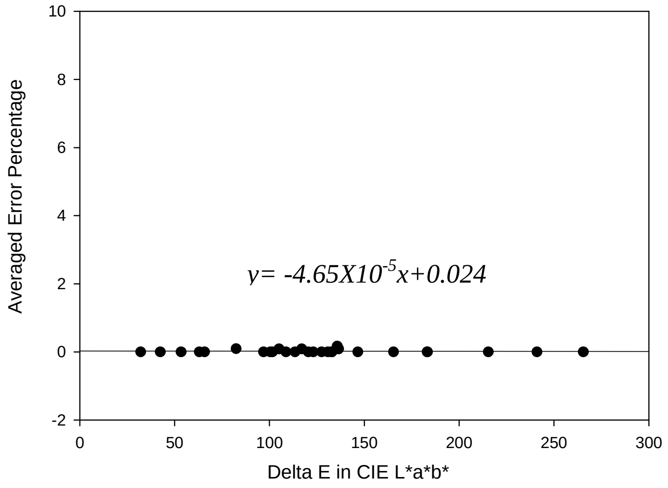


*Figure 7.1a. Second part VDT: Colour Normals-Delta E (L*a*b*) vs Averaged Error Percentage.*

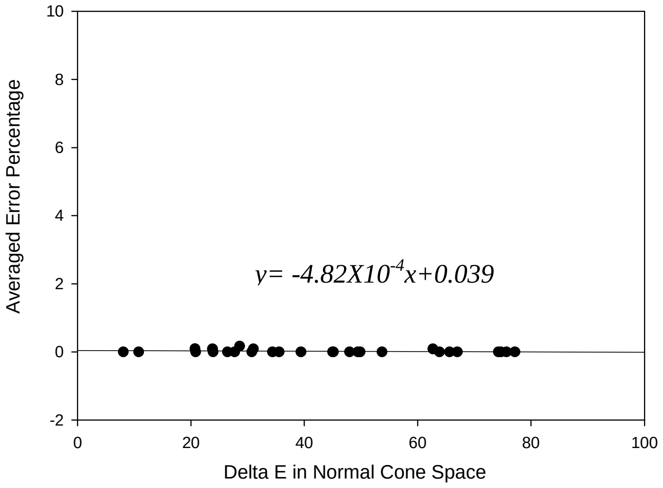


Figure 7.1b. Second part VDT: Colour Normals-Delta E (Normal Cone Space) vs Averaged Error Percentage

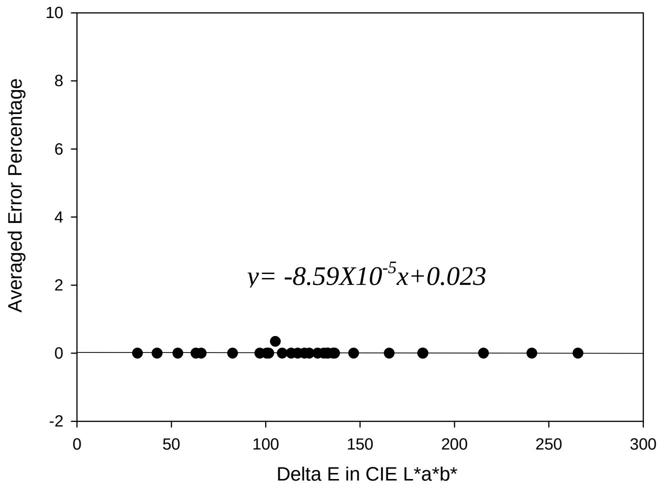


*Figure 7.2a. Second part VDT: Passed Colour-defectives-Delta E (L*a*b*) vs Averaged Error Percentage*

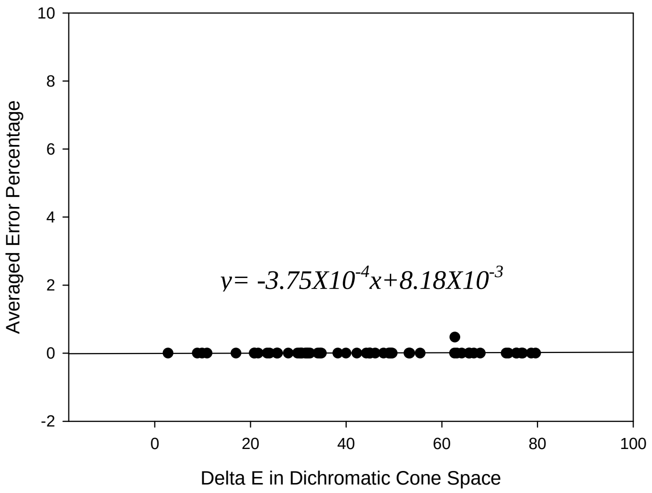


Figure 7.2b. Second part VDT: Passed Colour-defectives-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

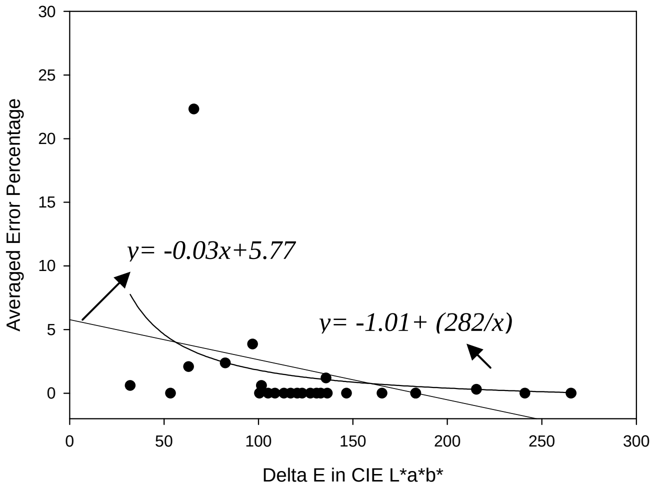


*Figure 7.3a. Second part VDT: Failed Colour-defectives-Delta E (L*a*b*) vs Averaged Error Percentage*

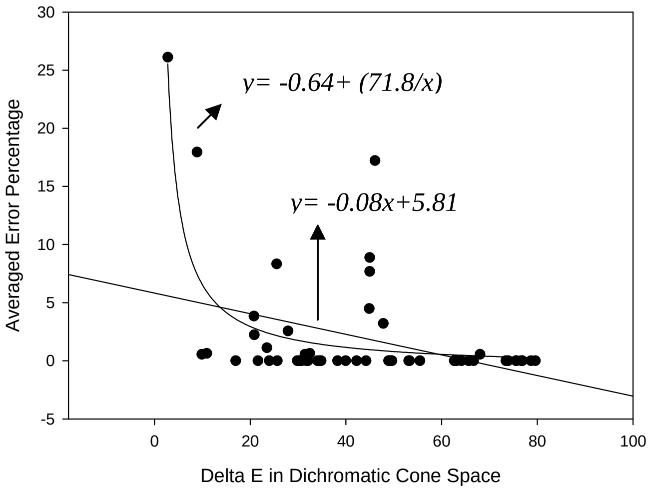


Figure 7.3b. Second part VDT: Failed Colour-defectives-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

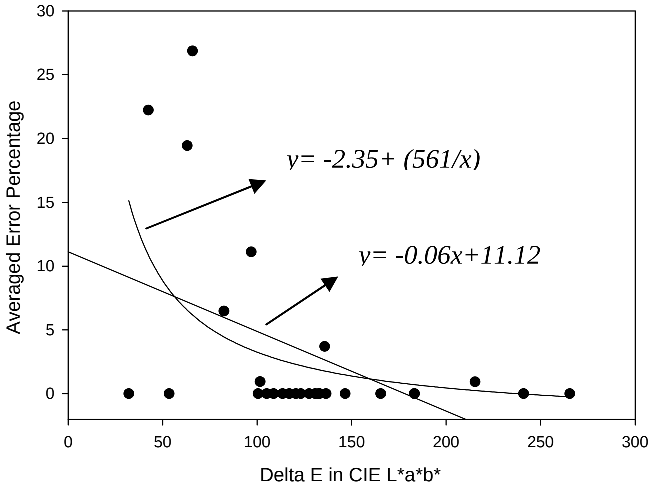


*Figure 7.4a. Second part VDT: Dichromats-Delta E (L*a*b*) vs Averaged Error Percentage*

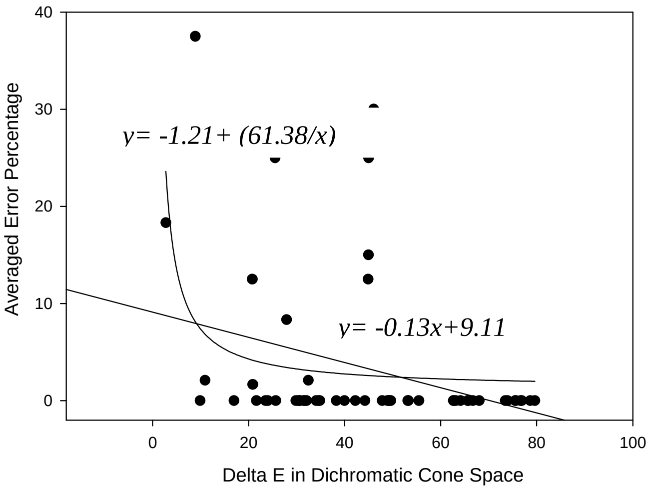


Figure 7.4b. Second part VDT: Dichromats-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

Correlation between averaged errors and colour difference in CIE Lab space (left side graphs) and normal and dichromatic cone space (right side graphs) for the third part(Authority) of the VDT test for colour normals, passed and failed colour defectives and dichromats.

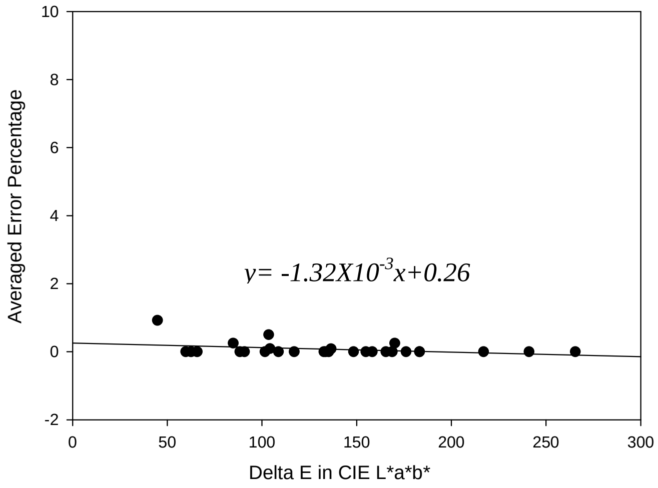


*Figure 7.5a. Third part VDT: Colour Normals-Delta E (L*a*b*) vs Averaged Error Percentage*

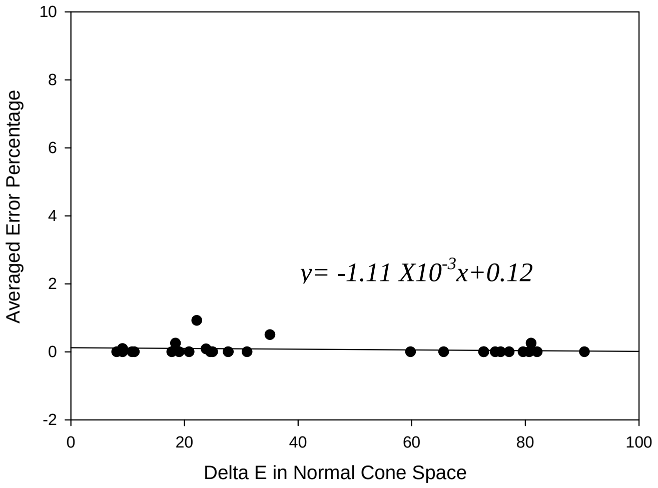


Figure 7.5b. Third part VDT: Colour Normals-Delta E (Normal Cone Space) vs Averaged Error Percentage

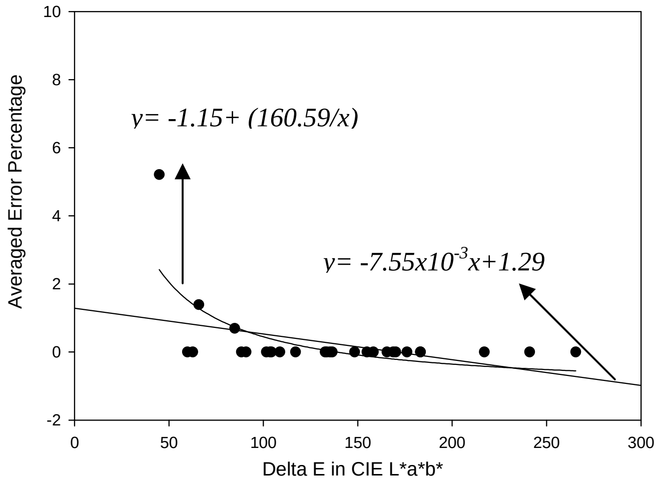


*Figure 7.6a. Third part VDT: Passed Colour-defectives-Delta E (L*a*b*) vs Averaged Error Percentage*

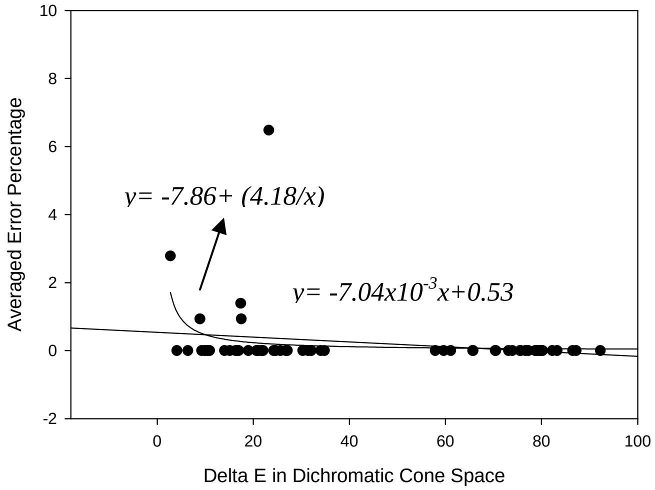


Figure 7.6b. Third part VDT: Passed Colour-defectives-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

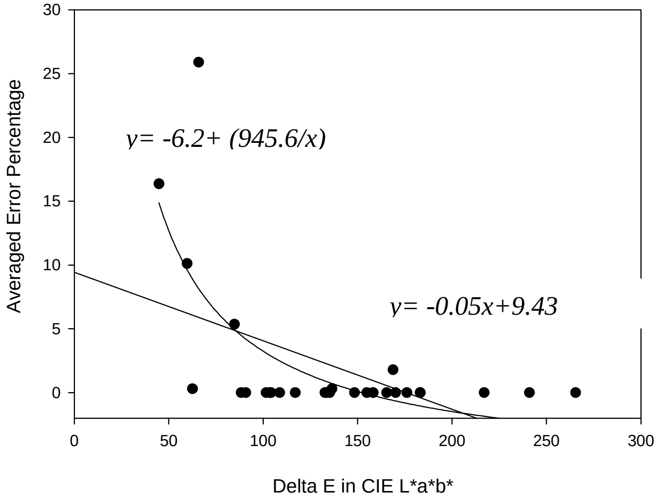


*Figure 7.7a. Third part VDT: Failed Colour-defectives-Delta E (L*a*b*) vs Averaged Error Percentage*

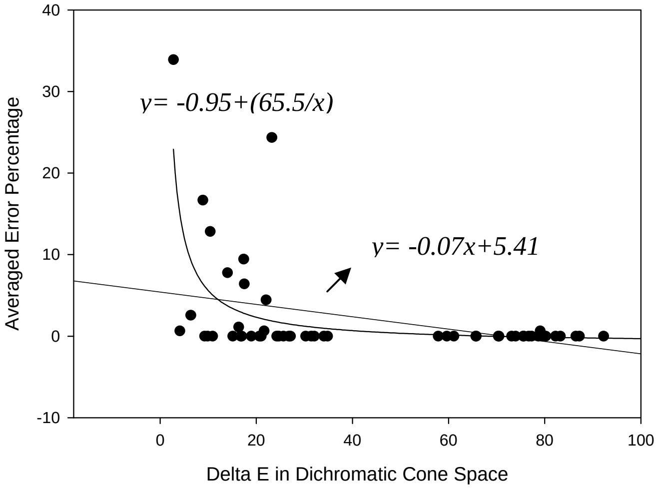


Figure 7.7b. Third part VDT: Failed Colour-defectives-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

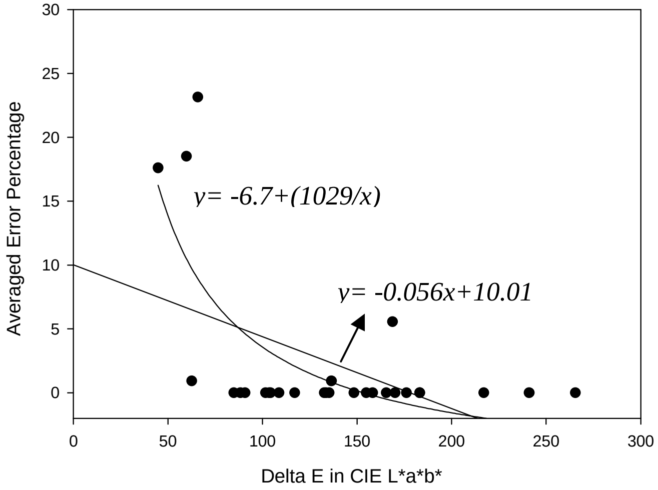


*Figure 7.8a. Third part VDT: Dichromats-Delta E (L*a*b*) vs Averaged Error Percentage*

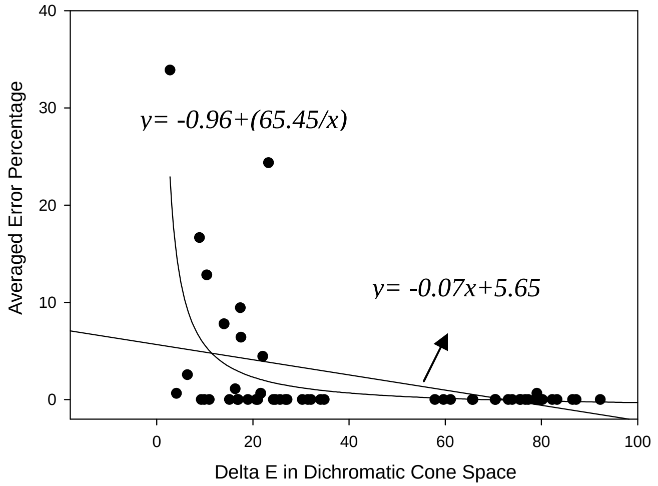


Figure 7.8b. Third part VDT: Dichromats-Delta E (Dichromatic Cone Space) vs Averaged Error Percentage

7.2. Discussion & Conclusion

The correlation results between the errors and colour-difference (ΔE*) indicate a weak inverse linear correlation between the mean errors and the colour-difference of the VDT colours in CIEL*a*b* and cone colour spaces. The best linear correlation was at 0.49 for the failed colour-defectives and dichromats and this improved to 0.70 with a nonlinear regression fit. One finding to note is that colour-normals were prone to make errors when the colour-difference (ΔE*=L*a*b*) was less than 60 units and colour-defectives made errors when the colour-difference was less than 150 units. The superior nonlinear fit indicates that the errors decrease rapidly once the colour-difference exceeds 60 (for colour-defectives). However, a colour difference less then these values does not guarantee that an error would occur. There were numerous colours, where the colour difference (ΔE*) was quite small and the error rate was zero.

One possible reason for the weak correlations is that the colour difference does not segregate luminance differences from hue differences. Thus a small ΔE* could be due to purely luminance difference that is easily discriminable. For example, the ΔE* for grey and white is 53 and there are almost no grey-white confusions. Similarly the ΔE* for white and blue-green was also small (32) and there were no mistakes. Nevertheless, a multiple regression performed on the data (Appendix 7.2) which treated the luminance and hue differences as independent variables did not improve the correlation. This last result may be due to the fact that a sufficient luminance difference did not necessarily guarantee that two colours on the same line of confusion will be discriminable. For example, ΔE* for the grey and blue-green is 44, however, 25% of the total errors occurred in this combination. As mentioned in the previous chapter, this may indicate an interaction with the average luminance of a pair of colours that may be confused or an interaction with the surrounding stimuli which also vary in colour and luminance. Finally, the ΔE values are related to threshold difference in colour. It is possible that certain colours do appear different to the colour-defective, but the ΔE is not large enough to identify the colours with two different categorical hues. For example, the blue-green may look slightly different from grey, but not sufficiently different so that person identifies it as blue-green.

The correlation for the colour-defective results was similar whether the colour difference (ΔE*) was calculated in colour-normal or dichromatic space. This result suggest that it is unnecessary to use abnormal colour spaces to predict error rates on display colours at least for these colour sets, although the transformation into dichromatic space does provide the colour-normal an appreciation of the problems individuals with severe colour vision deficiencies have in distinguish colours.

In summary, the errors were more frequent between colours that lie near the same line of confusion and had a similar luminance. Nevertheless, the interaction between luminance and location with respect to the lines of confusion was not easy to interpret. Examining where the colours fall with respect to the lines of confusion and the luminance differences does provide a very crude rule that if the colours are near the same line of confusion, the colour normal luminance ratio is less than 4 and if the colour-difference (ΔE*=L*a*b*) between the confusion colours is less than 60 units (in normal colour space), then errors could occur. If the colour normal luminance ratio is greater than 4 then errors, and if the colour-difference between the confusion colours exceeds 60 units in normal colour space (e.g. red-green, orange-green, red-yellow and orange-yellow) then the errors are unlikely or less frequent.

In addition to the making more errors on the VDT test, colour defectives typically take more time in making colour-related judgments. The next chapter presents the time taken by the colour defectives to complete the test and the results.

References

1. Wyszecki G, S. W. (1982). "Colour Science: Concepts and Methods, Quantitative Data and Formulae." Second Edition.
2. Golz, J. and D. I. MacLeod (2003). "Colourimetry for CRT displays." J Opt Soc Am A Opt Image Sci Vis. **20**(5): 769-81.
3. Brettel, H., F. Vienot, et al. (1997). "Computerized simulation of colour appearance for dichromats." J Opt Soc Am A Opt Image Sci Vis. **14**(10): 2647-55.

CHAPTER 8

Results IV. Time Taken to Complete the VDT Test

This chapter compares the time taken by the colour-normals and colour-defective subjects to complete the VDT test. This chapter has been accepted for publication in Optometry and Vision Science. Hence the term colour has been mentioned as color (American English) and the colour-defectives are mentioned as color-deficients.

8.1. Introduction

The advent of visual display terminals (VDTs) has increased the use of color to convey information in a variety of occupations. While the increased use of colors codes has often helped improve the flow of information for individuals with normal color vision (color-normals), it has often proven to be problematic for individuals with congenital red-green color vision deficiencies (color-deficient). Not surprisingly, the problems that color-deficients experience in identifying signal light colors,[1, 2, 14] resistor colors,[3] and wire colors[4] have also been shown for information displayed on VDTs.[5-15]

In addition to making more errors on VDT color displays, color-deficients were also slower in making their responses. Table 8.1 summarizes several studies on the response times for color-deficients to recognize individual colors displayed on a VDT and individual colored signal lights. The summary shows that as a group color-deficients were appreciably slower in recognizing individual VDT colors. Their response times were 10% to 52% longer than color-normals.[5, 9, 10, 15] The increase in response times was generally greatest for the deuteranopes followed by the protanopes, deuteranomals and perhaps protanomals.[5] The one exception was protanopes had longer response times (~ 30%) in identifying a red warning display.[9] The number of protanomals has been small in these studies so it is more difficult to determine their relative performance. Of particular interest is Olson and Brewer's[10] study which used two color schemes for displaying geographic maps on the VDTs. One scheme was designed to accommodate red-green color-deficients by minimizing color-confusions and the other design was to include as many potentially confusing colors for red-green color-deficients as possible. The response times of the color-deficients were 52% longer than color-normals on the potentially confusing map and 22% longer on the maps designed to minimize color-confusions. The increase in response times for the minimum confusion color set was similar to Cole and Macdonald's[9] result in that the color-deficient group was also slower to respond even to black-white photographs compared to color-normals. However, Cole and Macdonald attributed the slower response times of the color-deficients to their higher age and lower level of education.

In a more recent study, Cole et al.[15] found similar results to Olson and Brewer in that color-deficients can be slower to respond to when color is used in a display regardless of whether, or not, it is used to code information. They measured search times for an object that could be distinguished from the distracters by shape, color, or both. As expected, color-deficients had longer search times when the object's color was typically confused with one, or more of the distracter colors. However, they also reported that the search times for some color-deficients were prolonged for displays in which the color of the object and distracters were both green, yellow, or red.

The response times for identifying signal light colors showed a similar trend in the response times. Color-deficients had longer response times in general with dichromats having the longest times and protans having longer response times for red lights. However, the increase in response times relative to color-normals was often greater for the signal light task compared to the VDT displays, especially for the dichromats. In some cases, the dichromatic response times were nearly twice as long as color-normals. Two possible reasons for the relatively longer response times for the signal lights were that the signal lights were much smaller in angular size and their luminous contrasts were probably closer to the color-deficients threshold. Both factors would make colors more difficult to identify.

Table 8.1. Summary of studies measuring reaction and response times for identifying colors by color-normals and color-deficients.

Study	Color Task	Subjects		Increase in Time taken by Color-defectives compared to color-normals
		Color-Normals	Color- Deficients	
Nathan et al.[1]	Response times to R, G, Y signals lights	6	29 Deuteranomals: 14 Protanomals: 2 Deuteranopes: 6 Protanopes: 7	Range: 42 to 98% 52% 42% 70% 98%
Cole & Brown[2]	Response times to R, G, Y signals lights	11	Protanopes: 8	~400%
Bergman & Duijnhouwer[5]	Response times to VDT color codes (R, G, Y, C, B, M, W)	28	76 Deuteranomals: 16 Protanomals: 4 Deuteranopes: 28 Protanopes: 28	Range: 16 to 42% 16% NA 42% 24%
Cole & Macdonald[9]	Response times to EFIS-Video color coded display	12	18 Deuteranomals: 6 Deuteranopes: 6 Protanopes: 6	Range: 0 to 43% Deuteranomals Range:1-43% Deuteranopes Range: 6-31% Protanopes Range: 0-35%
Olson & Brewer[10]	Response times to color-coded maps displayed in VDT a)Accommodating rendition b)Color-confusing rendition	32	32	a) 22% b) 52%
Atchison et al.[14]	Reaction to Simulated Traffic signal (R, G, Y lights)	20	49 Deutans: 25 Protans: 24	Range: 35-85% For red signal: Deuteranopes: 53% For red signal: Protanopes: 35% For yellow signal: Deuteranopes: 85% For yellow signal: Protanopes: 53% Green no difference
Cole et al.[15]	Visual search times to identify diamond shaped target from distractive color and/or shapes in VDT display. The tasks are summarized as T1=target identified by shape and color differences and T2= target identified by shape only. (R, G, Y, B, W)	6	29 Deuteranomals: 9 Protanomals: 7 Deuteranopes: 6 Protanopes: 7	Range for Tasks T1, T2 Deuteranomals: 0-43%, 0% Protanomals: 0-32%, 0% Deuteranopes: 0-53%,0-35% Protanopes: 0-64 %, 0-22%
Current Study	Time to complete identification tasks for VDT colors (R, Y, G, DG, B, W, GY, O, P, BG)	81	44 Deuteranomals: 21 Protanomals: 12 Deuteranopes: 3 Protanopes: 5	Range: 10-34% Deuteranomals: 28% Protanomals: 15% Deuteranopes: 25% Protanopes: 10%

VDT= Video Display Terminal; EFIS= Electronic Flight Information System; R= Red, G=Green, Y=Yellow, O=Orange, W= White, B= Blue, BG= Blue-Green, DG=Dark-Green, GY= Gray, P=Purple, M=Magenta, C= Cyan.

Most studies which used color coded video generated display as their tasks, informed the subjects that they were either timed or limited in the amount of time they had to make individual responses.[5, 9, 10] There is little information as to how color-deficients perform when there are no imposed time constraints and subjects are allowed to proceed at their own pace for an extended period of time. The objectives of this study were to determine 1) how much longer color-deficients take to complete a VDT based color naming test which typically takes 8 to 10 minutes to complete without imposed time constraints; 2) whether color-deficients with an error rate within normal limits require a similar amount of time to complete the task; 3) whether familiarity with task has a differential effect on the time to complete the task for color-normals and color-deficients; 4) whether there is any correlation between errors and time to complete the task; and 5) whether there was a correlation between the Nagel anomaloscope range and Farnsworth D-15 with the time to complete the VDT test.

8.2. Methods

8.2.1. The Color Naming Test

The VDT test was developed as an occupational color vision test for Canadian Pacific railway dispatchers. The railway dispatchers monitor and control train movement using a VDT display. These displays use extensive non redundant color information to relay information on the track status, signal status, and rules that govern the track usage to the railway dispatcher. The actual test used the same computer, monitor, and graphic card as the dispatch centre. The test has been described in detail previously.[13] Briefly, the VDT test was comprised of three parts. The first part evaluated a person’s ability to identify 20 equilateral triangles as either red or green. The angular subtense of the triangle's side was 14 min of arc. There were 10 of each color displayed in a single column on the monitor. The position of the colors within the column was determined randomly.

The second part evaluated subjects’ ability to identify a set of colored rectangle icons as red, yellow, green, blue-green, dark-blue, purple, white or grey. The third part evaluated their ability to identify another set of rectangle icons as red, yellow, orange, green, dark green, light blue, dark blue or grey. The different color sets were based on two separate color codes used in the dispatch center. Table 8.2 lists the chromaticity coordinates and luminances of individual colors. The colors were measured with Minolta CS-100 Chroma meter with a 10x close-up lens. All the colors were displayed within a black background with a luminance less than 0.10 cd/m^2.

Table 8.2. Mean chromaticity coordinates and luminances of the colors used in the VDT test.

Color	**Chromaticity Coordinates**		**Luminance (cd/m^2)**
	x	**y**	
Red	0.616	0.355	11.86
Orange	0.581	0.387	20.13
Yellow	0.410	0.518	63.28
Green	0.290	0.604	43.86
White	0.292	0.304	65.77
Dark Green	0.237	0.391	9.99
Grey	0.280	0.301	11.88
Blue	0.145	0.068	8.39
Purple	0.276	0.147	12.33
Light Blue	0.197	0.226	36.32
Blue Green	0.206	0.288	26.02

In parts 2 and 3, the colored rectangles were arranged in a 2 by 8 array on the screen with each color being presented twice. The position within the array was determined randomly. Three different arrays were presented in both parts 2 and 3 so that each color was presented a total of 6 times. The angular size of rectangle was 35 by 11 min of arc. The test program recorded the error score and the total time to complete the task. The time to complete included reading the instructions, viewing examples, and addressing any questions that may arise about the procedure or test.

8.2.2. Subjects

Data were obtained from a subset of the subjects reported in the previous study.[13] In order to be included in this analysis, subjects had to perform the test twice and either pass the test at both sessions or fail the test at both sessions. In addition, all subjects had to have a minimum binocular visual acuity of 6/9 (20/30) for distance and near (i.e. 40 cm) with no known ocular pathology. There were 81 color-normals and 41 color vision deficient's (21 deuteranomalous, 12 protanomalous, 3 deuteranopes, and 5 protanopes) who met the inclusion criteria. The anomalous trichromats included three extreme deuteranomalous subjects. The larger number of color-normal subjects was selected to ensure that we had a large range of young adult color-normal responses and completion times. The mean age of the color-normals and the color- deficients was 30 years. Color vision was classified using the Nagel anomaloscope in the white adaptation mode.[16] The mean matching range for color-normal's and anomalous trichromats was 2.5 (± 1, standard deviation) and 12 (± 13, standard deviation) units respectively. There were 3 extreme anomalous trichromats.

The subjects gave informed written consent before participating and the study was approved by University of Waterloo's Office of Research Ethics.

In addition to the VDT test, subjects were also administered a variety of screening tests and the Farnsworth D-15. For the D-15 test, the subjects were asked to arrange the caps in the box by finding the moveable cap that was most similar in color to the last one placed in the box. The test results were evaluated by visual inspection and the number of crossings was noted. (A major crossing is defined as a difference between adjacent caps which is greater than 2).

8.2.3. Procedure

Each part of the VDT test was preceded by written instructions and examples of all the colors in the set. Subjects were free to ask any questions during the test. They entered their responses by clicking the color name below the colored icon. They could change any of their responses while a given array was displayed, but they could not go back to any of the previous arrays. Although the time to complete the test was recorded, the subjects were told that they could take as long as necessary. They were not informed that they were going to be timed at either session. There were two sessions separated by approximately two weeks.

Pass/fail was based on the 99th percentile color-normal total error score for each session. This corresponds to a failure when more than two errors were made on the first session and 1 error on the second session. In both sessions, misnaming orange as red was not counted as a mistake because 4 % percent (averaged across sessions) of the color-normals made this mistake multiple times within a session.

8.3. Results

Based on the failure criterion for each session, 49% of the color-deficients (20 subjects) passed the test at both sessions and the remaining 51% failed (21 subjects) the test at both sessions. All the dichromats failed both the sessions.

Figure 8.1 shows the time taken to complete the test at each session by color-normals, color-deficients who passed and color-deficients who failed. The first result to note is that the time to complete the task was approximately 18% shorter at the second session for all three groups. Repeated measures ANOVA showed that the session effect was significant (F=35.2, df=1, 119, p<0.05), but there was no significant interaction (F=1.88, df=2, 119, p>0.05) between session and group indicating that the reduction in the completion time was similar for all three groups. Figure 8.1 also shows that, although there was substantial overlap in the times to complete, color-deficients took more time relative to color-normals. Repeated ANOVA showed a significant group effect (F=16.67, df=2, 119, p<0.05).

Because there was no significant interaction between the groups and session, the remaining analyses will be based on the completion times averaged across the two sessions. These values showed that the color-deficients who passed the test took 13.5% longer than color-normals and color-deficients who failed took 29.5% longer. ANOVA (F=13.67, df=2, 119, p<0.001) and post hoc (Tukey HSD p<0.05) on mean values showed that the color-normals were significantly faster than color-deficients who passed and failed the test, but the two color vision deficient's groups were not significantly different from each other. Table 8.3 summarizes the results of the different categories of color-deficients and color-normals. Comparisons between the protans and deutans within each sub classification did not reveal any significant differences.

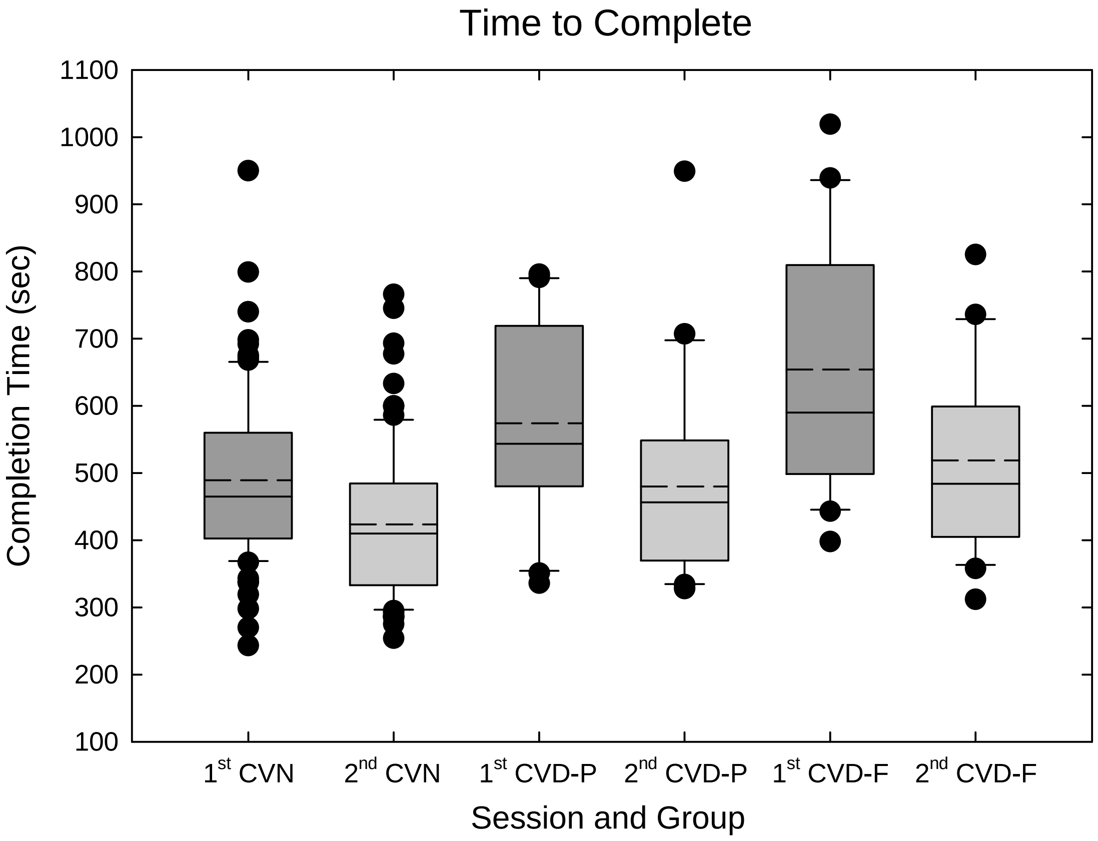


Figure 8.1. Box plots of the completion times for Color-normals (First session-1st CVN, Second session-2nd CVN), Color-deficients who passed (First session-1st CVD-P, Second session-2nd CVD-P) and Color-deficients who failed the VDT test (First session-1st CVD-F, Second session-2nd CVD-F). The error bars represents the 5th and 95th confidence intervals. The top and bottom of the box indicates the 75th and 25th percentile. The solid line inside the box is the median and the dashed dotted line is the mean.

Table 8.3. Mean times taken by different subject groups to perform the test across both the sessions and the percentage difference from the color-normal's times.

Groups (sample size)	**Mean Time to complete in seconds (± SD)**	**Completion time relative to color-normals (% longer)**
Color-Normals (81)	457 (±117)	0
Color vision-deficients (41)	556 (±147)	22
All Protans (Passed+Failed) (17)	532 (±157)	16
All Deutans (Passed+Failed) (27)	575 (±161)	26
Color vision-deficients who passed (20)	522 (±147)	14
Protanomalous (6)	460 (±122)	0.01
Deuteranomalous (14)	547 (±155)	20
Color vision-deficients who failed (21)	588 (±161)	29
Dichromats (8)	528 (±115)	15
Protanopes (5)	503 (±80)	10
Deuteranopes (3)	568 (±172)	24
Anomalous Trichromats (13)	621 (±174)	36
Protanomalous (6)	635 (±168)	39
Deuteranomalous (7)	627 (±188)	37

SD= Standard Deviation

Correlations with time to complete.

Figures 8.2 and 8.3 show the relationships between the mean completion time and the VDT errors (averaged across both the sessions). The subjects were divided into protans and deutans who passed and those who failed. There was no significant relationship between time to complete the task and errors for either group of the protan and deutan subjects or the group as a whole. There were no significant correlations between the Nagel Anomaloscope range (r = -0.09, p>0.05) or D-15 (r = 0.20, p>0.05) and time to complete for the color-deficient group.

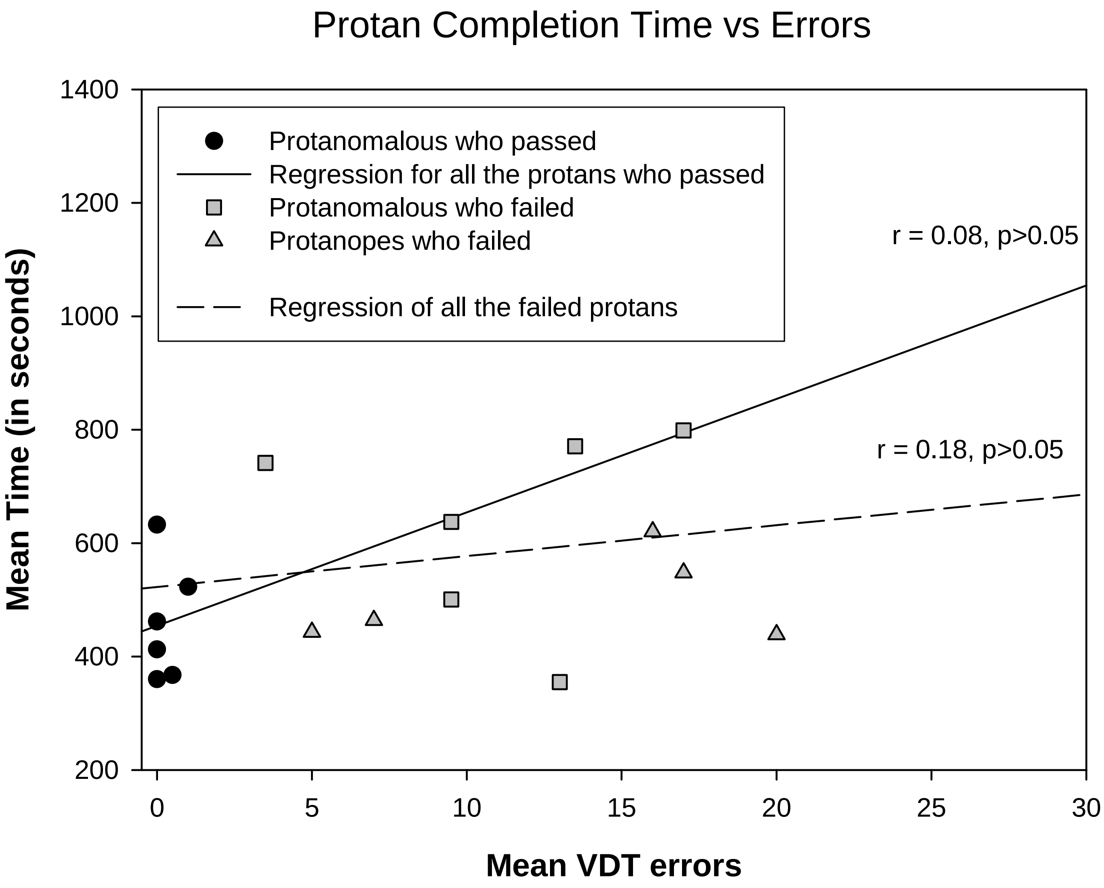


Figure 8.2. VDT task completion time and the number of errors for the protans averaged across both sessions. The *r* values are the Pearson correlation coefficient.

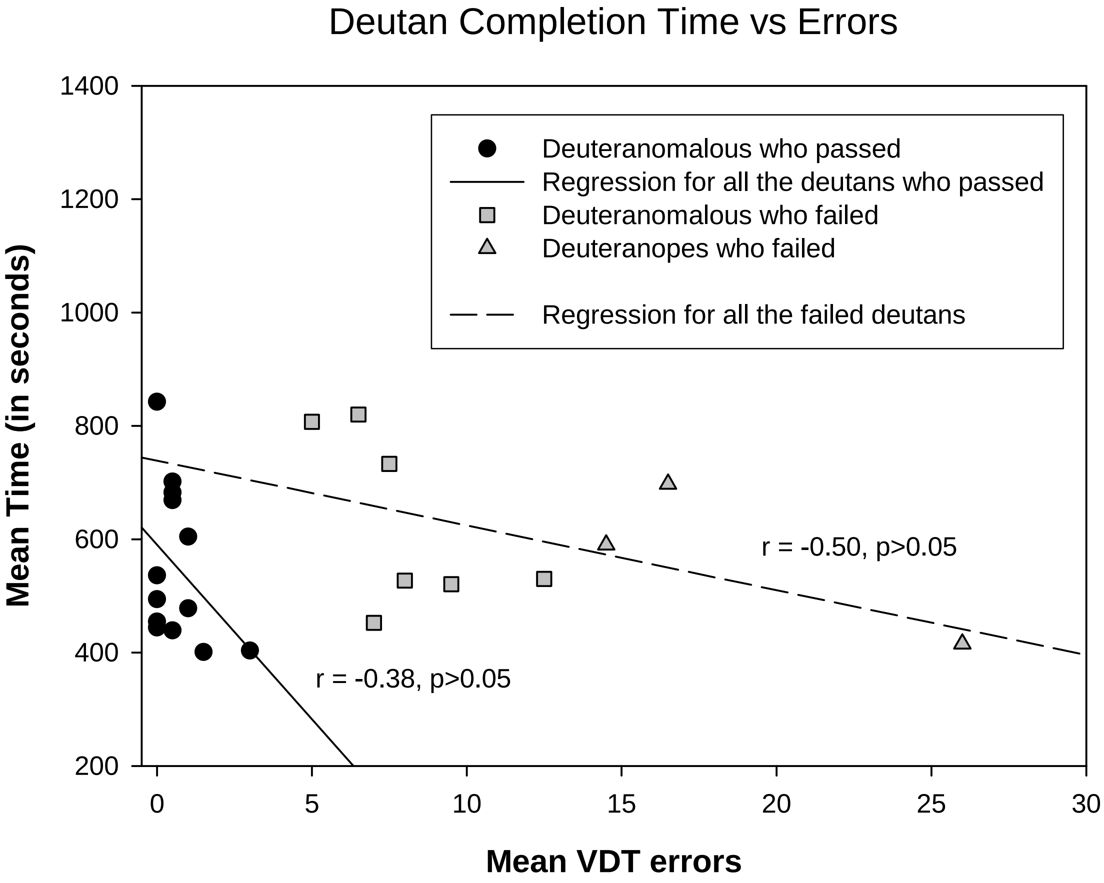


Figure 8.3. VDT task completion time and the number of errors for the deutans averaged across both sessions. The *r* values are the Pearson correlation coefficient.

8.4. Discussion

Regardless of whether they passed or failed the VDT test, the color-deficients took approximately 22% longer to complete a color-identification task than color-normals. These results were consistent with the previous findings of color-deficients having longer response times to identify individual VDT colors.[5, 9, 10] It is interesting that the average increase in response times to individual stimuli were reflected in an overall increase in the time to complete a series of color identifying tasks which took approximately 8 to10 minutes to complete.

Figure 8.1 shows that, despite the slower completion times for the color-deficients, there was a fair degree of overlap in the distributions. Using the 99th percentile completion time of the color-normals as a cut-off value, only 4 color-deficients took longer in the first session and only 3 individuals took longer in the second session. There was one color-deficient subject who took longer than the 99th percentile completion time of the color-normals in both sessions. These 7 subjects also failed the test at both sessions based their number of errors. All the color-deficients who passed the test completed the VDT test within the 99th percentile time for color-normals. Thus in terms of the number of errors and completion time, this group of color-deficients performed better than the worst color normal, although the mean completion time of the group was longer.

Subjects were informed that they could take as long as they desired to complete the test and so there was no time pressure for either the color-normals or color-deficients. One might expect that a number of the color-deficients would have taken an exorbitant amount of time to complete the task. This result did not occur frequently. This may have been partially due to the financial compensation for participating in the study. All subjects were compensated the same amount regardless of how long they took to complete the test and there was no reward or penalty based on the test error score. This may have been a disincentive to take one’s time in performing the task for some individuals. If that was the case, then one would have expected individuals with the faster completion times to have more errors. This trend was not evident in the data.

The total time to complete the VDT included instructions, examples and any questions which could have arisen during the test. One might anticipate that the color-deficient group would have required more time to complete the test based on the extra time required to familiarize themselves with the examples. Familiarization would be less of an issue at the second session and so they would have shown a larger decrease in completion times relative to color-normals. Although there was a significant reduction in completion times for all groups, there was no significant interaction between the groups and session which indicates that familiarization with the test and practice did not differentially affect the color- deficients' completion times.

The fact that there was no time constraint may be the reason for the general lack of a correlation between the number of errors and the completion time for the color vision deficient's groups. There were also no significant differences between protans and deutans in terms of the completion times. Nevertheless the deutans did show some trends that were reported by others.[14] This was longer time to complete compared to the protans. This trend may have been partially due to the fact that there were more potential confusions in this study for the deutans compared to the protans.

The increase in completion times for the VDT test was similar to the increase in response times to VDT colors, but often less than the increment found for signal lights (See Table

8.1). There are several possible reasons for the difference in response times for signal lights and VDT colors. First, the signal lights were more difficult to identify because of their small angular size and the brightness contrast was often closer to the subjects' threshold, particularly the red lights for the protan observers. In addition to having stimuli that were easier to identify in the VDT tasks, the larger, brighter, objects may have made detection of brightness differences between colors more obvious which may have aided individuals who use brightness information in color identification. It should be noted that the luminances of the colored icons was not varied in this study so that brightness information could be used as a secondary clue. We elected not to randomize the luminance because luminance measurements of the actual dispatch monitors showed that, although the average luminances of the display colors varied, the relative difference between colors remained the essentially the same. This result indicated that brightness information could be used as a reliable secondary clue in this case.

Another reason why the response times were longer for signal lights colors could be that relative number of potential confusions was higher for the lights. That is, of the 3 colored signal lights usually presented in the signal light studies, at least 2 of the 3 lights were on the same line of confusion. In the VDT task, one color could only be confused with either 2 or 3 of the seven other colors which would have made the overall task somewhat easier. Inclusion of a number of colors which are not likely to be confused has been shown to lower the average response times, although the times would not necessarily be as good as color-normals.[10]

8.5. Conclusions

1) Despite the lack of time constraints to complete an extended color naming task, our results were similar to previous findings in that the color-deficients required more 22% more time to complete a color identification task compared to color-normals. Color-deficients, especially the anomalous trichromats, who failed the test took the longest. 2) All the color-deficients who passed the VDT completed the test within the 99th percentile time for color-normals. Thus in terms of the number of errors and completion time, this group of color-deficients performed better than the worst color normal, although the mean completion time of the group was longer. 3) Practice did reduce the time to complete the test, but the relative decrease in completion times was similar for both the color-deficients and color-normals indicating that all groups benefitted equally from the practice. 4) Although there was no significant correlation between the number of errors and time to complete the test, deutans did show an inverse relationship between the time to complete the test and error rates, whereas, there was no relationship evident for the protans and 5) There were no significant correlations between the Nagel Anomaloscope range or the Farnsworth D-15 test and time to complete for the color-deficient group.

References

1. Nathan J, Henry GH, Cole BL. Recognition of colored road traffic light signals by normal and color-vision-defective observers. J Opt Soc Am 1964;54:1041-5.
2. Cole BL, Brown B. Optimum intensity of red road-traffic signal lights for normal and protanopic observers. J Opt Soc Am 1966;56:516-22.
3. Walraven PL, Leebeek HL. Recognition of color code by normals and color defectives at several illumination levels. An evaluation study of the H.R.R.plates. Am J Optom Arch Am Acad Optom 1960;37:82-92.
4. Hovis JK, Lu, B. and Neumann, P. The ability of colour vision tests to predict performance in identification of wire colour. Can J Optom 1994;55:219–228.
5. Bergman H, Duijnhouwer, F. Recognition of VDU presented colors by color defective observers. Proc Hum Factors Soc 1980:611–16.
6. Verriest G, Andrew I, Uvijls A. Study of visual performance on a multi-color VDU of color defective and normal trichromatic subjects. Bull Soc Belge Ophthalmol 1985;216:27-33.
7. Kuyk TK, Veres JG 3rd, Lahey MA, Clark DJ. Ability of deutan color defectives to perform simulated air traffic control tasks. Am J Optom Physiol Opt. 1987;64:2-10.
8. Macdonald WA, Cole BL. Evaluating the role of colour in a flight information cockpit display. Ergonomics 1988;31:13-37
9. Cole B, Macdonald W. Defective color vision can impede information acquisition from redundantly color-coded video displays. Ophthalmic Physiol Opt 1988;8:198-210.

10. Olsen J, Brewer C. An evaluation of color selections to accommodate map users with color vision impairments. Ann Assoc Am Cartogr 1997;87:103–34.
11. Mahon L, Jacobs R. Electronic Flight Information System displays and color defective observers. Clin Exp Optom 1991;74:196-203.
12. Mertens HW, Milburn NJ. Performance of color-dependent air traffic control tasks as a function of color vision deficiency. Aviat Space Environ Med 1996;67:919-28.
13. Ramaswamy S, Hovis JK. Ability of the D-15 panel tests and HRR pseudoisochromatic plates to predict performance in naming VDT colors. Vis Neurosci 2004;21:455-60.
14. Atchison DA, Pedersen CA, Dain SJ, Wood JM. Traffic signal color recognition is a problem for both protan and deutan color-vision deficients. Hum Factors 2003; 45:495-503.
15. Cole, B. L., Maddocks, J. D, Sharpe, K. Visual search and the conspicuity of coloured targets for colour vision normal and colour vision deficient observers. Clin Exp Optom 2004; **87**(4-5): 294-304.
16. Pokorny J, S. V., Verriest G, Pinckers AJLG eds. Congenital and Acquired Colour Vision Defects." Newyork: Grune and Stratton. 1979.

CHAPTER 9

Results V. Comparison with Clinical Tests

9.1. Comparison with Clinical Tests

The objective of this study was to determine whether clinical colour vision tests can predict the performance on the VDT dispatch test. The clinical tests that were compared in this study were Farnsworth D-15, Adams D-15, Ishihara pseudoisochromatic plate (38 plate edition, 1991), HRR pseudoisochromatic plates (3rd edition, Richmond Products, 1991) and CN Lantern test. The reasons for selecting these tests were outlined in Chapter 3.

9.2. Preliminary analysis

In the preliminary analysis, [1] the VDT dispatch test was compared with Farnsworth D-15, Adams D-15, and HRR diagnostic pseudoisochromatic plates. The main purpose of this study was to determine the scoring criterion for each test which produced the maximum agreement with the VDT test and to determine which of the three tests had the highest level of agreement with the VDT test. The kappa (κ) coefficient of agreement (agreement between two sets of observers/methods on nominal scale correcting for the chance agreement), [2] predictive pass and predictive fail rates were used as the indices.

9.2.1. Kappa (κ) coefficient of agreement

The general finding was that the κ values for both D-15 tests were significantly above zero (i.e. chance agreement) and they were similar across the various D-15 failure criteria (Figure 9.1).

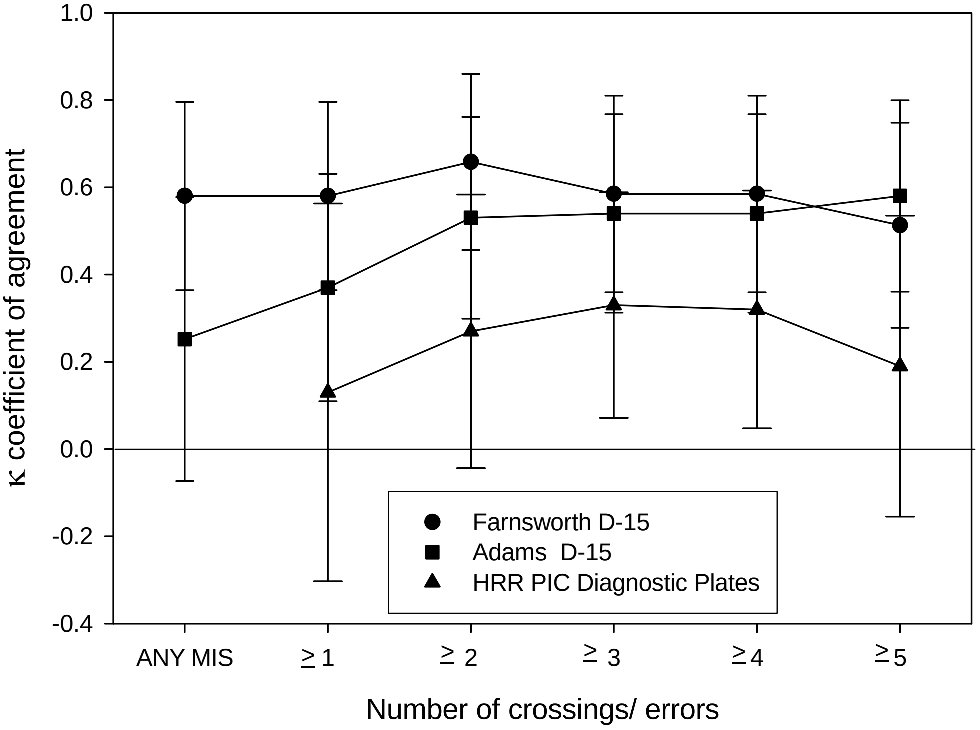


Figure 9.1. Coefficients of agreement between the VDT test and the clinical tests for different failure criteria. The numbers on the x-axis are various failure criteria based on the number of crossings on the D-15 tests or the number of errors on the HRR red-green diagnostic plates. ANY MIS means that any mistake (including transpositions) on the D-15 tests was a failure. (Figure reprinted with permission. Appendix 9.1)

The maximum κ for the Farnsworth D-15 and the Adams D-15 was 0.65 and 0.60 respectively. On the other hand, the level of agreement for the HRR diagnostic plates was lower across all failure criteria with only two criteria (≥3 and ≥4 errors) having a κvalue significantly above zero. (None of the colour-defective subjects made an error on the HRR blue-yellow diagnostic plates). However, even for these two criteria, the HRR plates had a significantly lower level of agreement than the Farnsworth D-15 based on the Farnsworth D-15 95% confidence intervals. Although all of the HRR κ values were lower than the Adams D-15, the difference did reach significance based on the 95% confidence intervals.

A side issue of this study was to determine whether there was good agreement between the Ishihara test and the Nagel anomaloscope using the current Railway Association of Canada guidelines [3] of more than 5 errors. There was 100% agreement between the Ishihara test using the Railway Association's failure criterion and anomaloscope as to who has normal and abnormal colour vision. The maximum number of errors made by a colour-normal was 1 and the minimum error made by colour-defective was 9. Our results were also in agreement with Birch's recommendation that more than 3 errors on the first 16 plates (ignoring responses on the hidden figures) is also an effective screening criterion.[4] All the colour-defectives made at least 4 errors on the transformation and vanishing plates.

9.2.2. Predictive Pass:

Predictive Pass: Figure 9.2 shows the predictive pass values (probability that a subject who passes the clinical test will pass the VDT test) for the three tests. The values for the

D-15 tests were similar to each other (for ≥2, ≥3, ≥4 crossings) and across the various scoring protocols. In contrast, the HRR predictive pass probabilities were more varied and were significantly lower than the Farnsworth D-15's when numerous errors are allowed on the HRR-3rd edition plates.

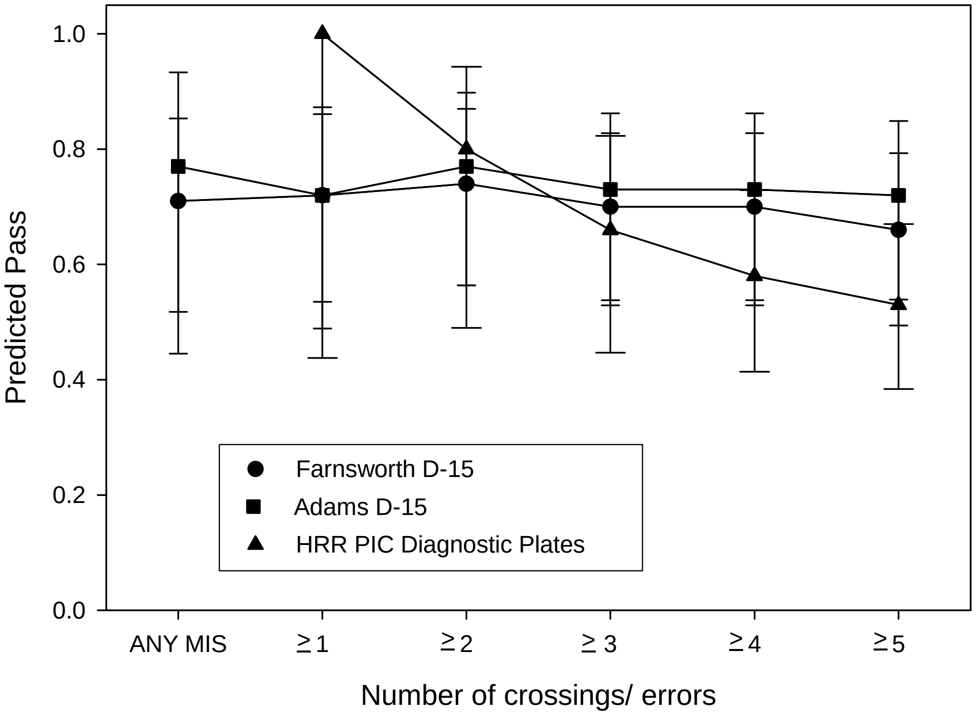


Figure 9.2. Predictive pass probabilities of the clinical tests for different failure criteria. The numbers on the x-axis are various failure criteria based on the number of crossings on the D-15 tests or the number of errors on the HRR red-green diagnostic plates. ANY MIS means that any mistake (including transpositions) on the D-15 tests was a failure. (Figure reprinted with permission)

9.2.3. Predictive Fail

Figure 9.3 shows the predictive fail values (probability that a subject who fails the clinical test will also fail the VDT test) for the tests. The values for the Farnsworth D-15 test were all above 0.5 and slightly higher than the Adams D-15 across all scoring protocols; however, they were not significantly different (based on 95% confidence intervals). In contrast, the HRR predictive fail probabilities were significantly lower than the Farnsworth D-15 and Adams D-15 for failure criteria between any mistake and >3 errors. The HRR test predictive failure rate increased (to match both D-15 rates) if the number of errors allowed on the diagnostic plates were increased to 4 or 5.

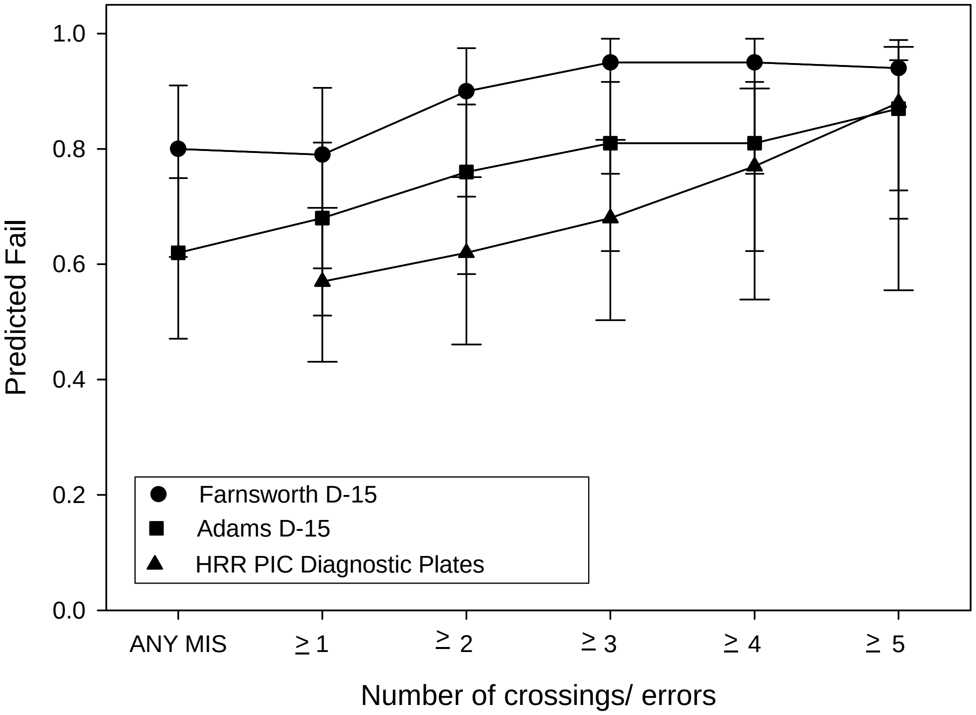


Figure 9.3. Predictive fail probabilities of the clinical tests for different failure criteria. The numbers on the x-axis are various failure criteria based on the number of crossings on the D-15 tests or the number of errors on the HRR red-green diagnostic plates. ANY MIS means that any mistake (including transpositions) on the D-15 tests was a failure. (Figure reprinted with permission)

9.2.4. Discussion

The primary aim of this study was to see whether the three clinical tests can adequately predict colour naming performance on a VDT based task. The preliminary analysis shows that all the three clinical tests had a significant level of agreement with the practical task. Whether the agreement is high enough to allow one to substitute the D-15 tests or HRR for the VDT test is open to discussion because the definition of an acceptable κ value varies across disciplines and safety concerns. However, a high level of agreement (e.g. > 0.95) between the clinical and VDT practical test along with a high predictive pass value for the clinical test (e. g. >0.95) would be preferred. None of the three tests met both these conditions.

For the D-15 tests, the predictive pass value was 0.70 across the failure criterion which indicates that about 70% of the colour-defectives who pass either of the D-15 tests will pass the VDT test. The results for the HRR suggest that a perfect score on the red-green diagnostic plates is highly predictive of a passing performance on the VDT test, but this conclusion was based on the small number (n=3) of subjects in this group. As more mistakes were allowed on the diagnostic plates, the predictive pass values became similar to, if not marginally lower, than the D-15 values.

The predictive fail values show that failing the Farnsworth D-15 was highly predictive of failing the VDT practical test in that over 90% of the people who failed the D-15 using multiple major crossings as a failure also failed the VDT test. The predictive values for the Adams D-15 were slightly lower and the HRR diagnostic plates were the least predictive of a failure. This trend in the predictive value is not surprising since subjects with the more severe defects were the ones most likely to fail both the D-15 and the VDT test.

Although the agreement between the Farnsworth D-15 and the VDT practical test was similar for all failure criteria, ≥2 crossings was preferred as a failure criterion because it is the standard failure criterion given in the instructions. We would not recommend using ≥3 crossings (more crossings) as a failure because one protanope would have passed the Farnsworth D-15 using this criterion and he did fail the VDT test. The same protanope passed the HRR plates with the ≥3 errors criterion. This latter finding that dichromats (especially protanopes) can make a small number of crossings and perhaps pass the Farnsworth D-15 is consistent with Birch's findings; however, none of the dichromats passed the Farnsworth D-15 using the standard scoring in this study. Birch recommended that any major crossing should be a failure on the Farnsworth D-15 to avoid passing dichromats. Although the agreement between any major crossing as a failure and the standard criterion of ≥2 major crossings was statistically identical, we would still prefer the standard failure because the predictive value for failing is marginally better than any major crossing without any reduction in the predictive value for passing . All dichromats failed the Adams D-15 using ≥3 major crossings as a failure.

Despite the less than ideal agreement with the VDT test, the Farnsworth D-15 test was a reasonable predictor of who will fail the VDT practical test. Ninety five percent who failed the D-15 using the Farnsworth criterion of $\geq$2 crossings also failed the VDT test. These results show that a failure on the Farnsworth D-15 is highly predictive of a failure on the VDT test so that any subject who fails the Farnsworth D-15 could be disqualified from working as a railway dispatcher with a false alarm rate varying between 5% using the standard scoring criterion. The HRR may also offer some efficiency in testing because of the high predictive pass value when no errors are allowed on the diagnostic red-green plates. However, because of the small number of subjects in this group (n=3), the predictive pass value is imprecise and it would be difficult to make any firm conclusions about using this criterion as a substitute for the VDT test.

9.3. Logistic Regression

Past studies [5-8] used a variety of indices to evaluate how well clinical tests predicted performance on colour related tasks. These indices include linear correlation, sensitivity and specificity, predictive values and kappa coefficients of agreement. Although these statistical indices may be useful when investigating how a single test compares, it may not be the most efficient method when a variety of clinical tests are included in the study. Examining individual test performance also limits the possibility of developing an efficient test battery.

To determine which clinical tests, either individually or in combination, can best predict the colour naming performance on the VDT test, logistic regression was used. Logistic regression was selected because it is best suited for situations where the dependent variable is dichotomous, when there is a nonlinear relationship between the dependant and the independent variables, and the distributions are non-normal distributions.[8]

Logistic regression [8] is a mathematical modeling approach used to predict the presence or absence of an outcome based on values of a set of predictor variables. This procedure builds a predictive model of group membership based on observed characteristics of each case. The analysis generates a function based on a linear combination of the predictor variables (clinical tests) that provide the best discrimination between the groups (e.g., pass/fail). Once the core model is established, model building was performed in several stages (stepwise analysis) based on significant predictors to arrive at the final model. In the stepwise analysis, all variables are reviewed and evaluated to determine which one(s) will contribute most to the discrimination between groups. At each step, the variable (clinical test) that contributes the least is removed and the process starts again with the remaining variables. The process stops once variables no longer improve the fit of the data using a given statistical criterion. This results in the minimum number of variables (clinical tests) that produce the optimum fit.

The clinical test parameters used in the Logistic regression (SPSS version 16) were the number of crossings on both D-15 tests, the range of acceptable matches on the Nagel anomaloscope, the number of errors on CN Lantern (4.6 m distance and 2.3 m distance), and the number of errors on the red-green diagnostic plates for both Ishihara and HRR pseudoisochromatic plates.[9] Both test distances for the CN Lantern were included in the analysis to determine whether either test distance could be a good predictor of the performance in the practical test.

In the current analysis, a Backward Stepwise Likelihood ratio method was used. In this method, once the core model is established, SPSS builds the equation starting with all variables and then removes them one by one if they do not contribute (significant) to the regression equation. The backward stepwise procedure was preferred over the forward stepwise because the backward procedure is less likely to exclude variables which provide a small but important contribution. Although both forward and backward stepwise techniques often generate identical results, the backward stepwise is more likely to uncover these relationships since all variables are initially included in the model.[10] Variables which contribute to the majority of the model can suppress (or mask) these smaller effects in forward procedures. The complete output results of the logistic regression model are displayed in Appendix 9.2 and the key results will be discussed here.

The classification table 9.1 compares the predictive results of using the logistic regression model with the actual data. For each case, the predicted response is “Fail” if that case’s model-predicted probability is greater than the cut-off value (in this case, the default of 0.5) and “Pass” if the model-predicted probability is less than 0.5. The percent correct at each step of the backwards regression shows how well the various models fit the data.

Table 9.1. Classification Table

		Step 1 Includes all Tests (Adams D-15 Farnsworth D-15 HRR, Ishihara, CN Lantern 4.6m CN Lantern 2.3m Anomaloscope Range)		**Step 2** CN Lantern 4.6 m Removed Includes Adams D-15 Farnsworth D-15 HRR, Ishihara, CN Lantern 2.3m Anomaloscope Range	**Step 3** Adams D-15 Removed Includes Farnsworth D-15 HRR, Ishihara, CN Lantern 2.3m Nagel Anomaloscope			**Step 4** CN Lantern 2.3 Removed Includes Farnsworth D-15 HRR, Ishihara, Nagel Anomaloscope			**Step 5** Ishihara Removed Includes Farnsworth D-15 HRR, Nagel Anomaloscope			**Step 6** HRR Removed Includes Farnsworth D-15 Nagel Anomaloscope		
	Observed	VDT test			VDT test			VDT test			VDT test			VDT test		
		Pass	Fail	Overall Percentage	Pass	Fail	Overall Percentage	Pass	Fail	Overall Percentage	Pass	Fail	Overall Percentage	Pass	Fail	Overall Percentage
Model Prediction	Pass	22	8	P.P=73.3	22	8	P.P=73.3	22	7	P.P=75.9	22	8	P.P=73.3	23	8	P.P=74.1
	Fail	2	20	P.F=90.9	2	20	P.F=90.9	2	21	P.F=91.3	2	20	P.F=90.9	1	20	P.F=95.2
	Percentage Correct	91.7*	71.4[#]	**80.8**	91.7	71.4	**80.8**	91.7	75	**82.7**	91.7	71.4	**80.8**	95.8	71.4	**82.7**

P.P=Predictive Pass; P.F=Predictive Failure

*Equals the test specificity

[#] Equals the test sensitivity

The general trend in the results was that the various versions of the model predicted approximately 82% of the cases correctly with simpler models providing slightly better predictions. Both step 4 and step 6 have the same maximum correct value of 82.7%. However, the table shows that there was a subtle difference between the two steps. The sensitivity of the model at step 4 was slightly higher than step 6, whereas the specificity at step 6 was slightly higher. This means that using just Farnsworth D-15 and Nagel Anomaloscope was a slightly better predictor of failure on the VDT test than the combination of Farnsworth D-15, HRR, Ishihara, and Nagel Anomaloscope in Step 4. However, these four tests were a slightly better predictor of who will pass the VDT test.

9.3.1. Discussion

Ideally, one would want to maximise the percent correct with a minimum number of predictor tests. Given the safety considerations, one would also want to have a test battery to have a high sensitivity. However, the logistic analysis shows that the variation in sensitivity and specificity across different models was marginal. The sensitivity varied from 71.4% to 75% and the specificity varied from 91.7% to 95.8%. The logistic analysis does indicate a test battery may help in refining a clinician's ability to counsel a patient. Step 6 indicates that both the anomaloscope range can be used along with the Farnsworth D-15 such that no errors on the Farnsworth D-15 and a small range on the anomaloscope allow the clinician to advise the patient that he has a good chance to pass the VDT test. Similarly, a large range on the anomaloscope and many errors on the Farnsworth D-15 would indicate that the person has a high probability of failing.

However, it is more difficult to predict when there are minimal errors (< 2 crossings) on the Farnsworth D-15 and the Nagel range indicates a mild to moderate defect.

Nevertheless, Table 9.2 shows that a logistic regression test battery defined in Step 6 gave results that were very similar to just the Farnsworth D-15 test in terms of agreement with the VDT test. There were, however, a couple of discrepancies between the predicted logistic model (step 6) tests and the Farnsworth D-15 alone. These were, one subject who failed the Farnsworth D-15 (and failed the VDT test) was predicted to pass based on the model; another subject who passed the D-15 (and passed the VDT) was predicted to fail by the model.

Table 9.2. Coefficient of agreement between the VDT test and the Farnsworth D-15 (≥ 2 crossing is a failure)

VDT test

		Pass	**Fail**	
Farnsworth D-15	**Pass**	**23**	**8**	**Predicted Pass= 0.74**
$\kappa = 0.658$ Se = 0.103	**Fail**	**1**	**20**	**Predicted Fail= 0.95**

(Model prediction: 82.7%) Specificity: 0.958 Sensitivity: 0.71

Despite the less-than-perfect prediction, both analyses show that the predictive value for failing the Farnsworth D-15 is very high and can be used by clinicians for counselling. The higher specificity and higher predictive value of the D-15 for failing a practical colour-naming task was generally different from the results reported by others. Table 9.3 shows that the sensitivity of the Farnsworth D-15 was reported to be higher than the specificity for a number of tasks. The difference may be the combination of the colours present in the practical test and the proportion of the different severities within the colour-defective sample.

Table 9.3. Sensitivity and specificity of D-15 test reported in few studies.

Study	**Task**	**Specificity (Pass both D-15 and Practical Task)**	**Sensitivity (Fail both D-15 & Practical Task)**	**Comments**
Kuyk et al [4]	Air Traffic Control. Both Surface colours and signal lights	0.58	0.95	Combined results for both protan and deutan subjects. Mild defects passed the D-15.
Hovis, et al [11]	Identifying wire colours	High voltage wires: 0.72 Low voltage wires: 0.93	High voltage wires: 0.73 Low voltage wires: 0.85	Low voltage wires contained more colours, were smaller in size and had more pastel shades
Mahon & Jacobs [12]	Colours on Video display (Electronic Flight Information System)	Single Colour Presentation: 0.5 Paired Colour Presentation: 0.58	Single Colour Presentation: 1.0 Paired Colour Presentation: 0.92	Values for the single presentation are based on the second series. The worst normal score was the cut off score for passing the practical.
Sui & Yap [5]	Road markings and single lights used in airports	0.44	0.88	Limited set of surface colours (n=3) and the colours and intensities of the signal lights were not given.
Cole & Orenstein [6]	Paint, thread, fabric samples	Large objects: 0.69 Small objects: 0.72 Combined: 0.74	Large objects: 0.81 Small objects: 0.74 Combined: 0.78	Subject age 11 to 65 yrs. Colours across materials were basically the same set, but may not have had the same brightness relationships

Although the agreement with the Farnsworth D-15 was good, it was not sensitive enough to replace the VDT practical test. The logistic regression analysis did not help solve this problem because none of the steps produced an appreciably better sensitivity. The sensitivity and predictive pass value of the clinical test battery can be improved to 0.9 using a moderate-to-severe classification (e.g. Nagel range $\geq$ 10 units or more than 1 error on the HRR diagnostic plates) to predict a failure on the VDT test. However, the specificity decreases to near 0.2.

The logistic regression analysis showed that any test battery using these clinical tests was no better than the Farnsworth D-15 alone in predicting who would pass or fail the VDT test. Neither the Farnsworth D-15 nor a test battery was sufficient to replace the practical test. Nevertheless, the results are useful for counselling purposes.

1. If a patient fails the Farnsworth D-15, then over 95% chance of failing the VDT test.
2. If a patient passes the Farnsworth D-15, then over 75% chance of passing the VDT test.
3. If a patient passes the Farnsworth D-15 and makes less than 2 errors on the HRR diagnostic plates or Nagel range less than 10 units, then they have 94% chance to pass the VDT test.

The results also imply that a test battery that allows a finer classification of the colour vision defect and assesses colour discrimination for colours other than red, green and yellow may be helpful in improving the sensitivity of the clinical assessment. These additional tests could include the Lanthony Desaturated D15 or the Farnsworth Munsell 100 Hue test.

References

1. Ramaswamy, S and Hovis JK (2004). Ability of the D-15 panel tests and HRR pseudoisochromatic plates to predict performance in naming VDT colours. Vis Neurosci **21**(3): 455-60.
2. Hays, W. (1981). Analysing Qualitative Data: Chi-Square Tests. Chapter 15. Statistics. New York, CBS College Publishing**:** 558-559.
3. Birch, J. (2001). "Diagnosis of defective colour vision, Second edition." Chapter 9: 108-109.
4. Railway Association of Canada (1990). “Canadian Rail Operating Rules. Montreal QC.” The Railway Association of Canada.
5. Kuyk, T.K., Veres, J.G., Lahey, M.A. & Clark, D.J. (1987). Ability of deutan colour defectives to perform simulated air traffic control tasks. American Journal of Optometry and Physiological Optics 64, 2–10.
6. Siu, A.W. & Yap, M. (2003). The performance of colour deficient individuals on airfield colour tasks. Aviation Space and Environmental Medicine 74, 546–550.
7. Cole, B. & Orenstein, J. (2003). Does the Farnsworth D-15 test predict the ability to name colours? Clinical and Experimental Optometry 86, 221–229.
8. Kleinbaum, D. G. (2002). "Logistic regression: a self-learning text." Statistics for biology and health: 513.
9. George, D. and P. Mallery (2001). Logisitic Regression. SPSS for Windows Step by Step: A Simple Guide and Reference,10.0 Update, (Toronto.: Allyn and Bacon, 2001). Chapter 25.

10. Menard, S. (2002). Applied logistic regression analysis, 2nd Edition. Thousand Oaks, CA: Sage Publications. Series: Quantitative Applications in the Social Sciences, No. 106. First ed., 1995.
11. Hovis JK, Lu, B. and Neumann, P (1994). The ability of colour vision tests to predict performance in identification of wire colour. Can. J. Optom. 1994;55: 219–227.
12. Mahon, L. and R. Jacobs (1991). "Electronic Flight Information System displays and color defective observers." Clinical and Experimental Optometry 74.6 Nov/Dec: 196-203.

CHAPTER 10

10.1. Conclusions

The purpose of this study was to establish pass/fail scores and repeatability of the VDT Dispatch colour vision test. The results indicate that the optimum criterion was to ignore any errors where orange was misnamed as red and then allow 2 errors on the first session and 1 error on the second session. Based on this criterion, 42% of the individuals with a colour vision defect (colour-defectives) could perform as well as colour-normals in identifying VDT railway display colours. The kappa coefficient of agreement for between the session repeatability was 0.85 for the colour-defective group. Nevertheless, there remains the issue as to whether individuals who make 1 to 2 errors should also fail the test if their errors are typical of colour-defectives. The difficulty in making this decision is that these errors are rarely repeatable when the number of mistakes is small.

The types of errors made by the colour-defectives who passed the VDT Dispatch test showed similar trends as the errors made by the colour-normals with one exception. A small number of these individuals (30%) occasionally identified green as yellow. However, this error is less critical than yellow called green errors.

The types and number of errors made by the colour-defectives who failed the Dispatch test were more pronounced and showed a trend that was consistent with their respective lines of confusion. All the dichromats failed the VDT test. Protans made about equal errors in the track status (part two) and the occupational authority (part three), whereas deutans made more errors on the occupational authority. This was primarily due to the two fold increase in errors on the third part by the deuteranomalous subjects.

As expected, the colour-defectives made more errors between colours that were near or along the same lines of confusions and the colours were nearly equal in luminance. Nevertheless, the correlation between errors made by colour-defectives who failed and the colour-difference (ΔE*) in CIE L*a*b* and cone colour spaces was only moderate at best. The lack of a strong correlation may have been due to an interaction between luminance differences, contrast effects from neighbouring stimuli, and the average luminance of a coloured-pair.   .

The time to complete the test for the colour-defectives (as a single group) was 22% longer than colour-normals. Colour-defectives who passed the test took 14% longer than colour-normals and colour defectives who failed took 30% longer. All groups showed a similar learning effect with an 18% reduction in mean times to complete the task at the second session. There was no significant correlation between the number of errors and time to complete or the clinical tests and completion times for any of the groups.

Another aim of the study was to examine the correlation between the VDT Dispatch test and the clinical colour vision tests. Logistic analysis results showed that the Farnsworth D-15 combined with the Nagel was the best predictor of the VDT Dispatch colour test performance. However, these results were similar to using the Farnsworth D-15 test alone. The correlations with the clinical tests revealed that the clinical tests could reasonably predict who was going to fail the Dispatch test, but could not predict accurately who was going to pass the Dispatch test. Ninety-five percent of the individuals who failed the Farnsworth D-15 or exceeded the criterion value on the combined D-15 and Nagel test battery also failed the Dispatch test. However, only 75% of the individuals who passed the Farnsworth D-15 or were below the criterion values for the combined D-15 and Nagel battery passed the VDT Dispatch colour test. Although the clinical tests were imperfect predictors of the Dispatch test outcome, the D-15 can be useful for counselling. Clinicians can advised their patients who fail the D-15 that they have a very high probability of failing the Dispatch test, whereas individuals who pass the D-15 have a reasonable probability of passing the Dispatch test, they still have a 25% chance of failing.

10.2. Future Directions

The first generation of the test was very basic. It may be worthwhile to redesign the test layout, so that it is more similar to actual railway traffic control display as shown in Figure 2.1. This new design would create more realistic displays for studying the effects and cognitive load and fatigue on colour identification. Two issues that should be investigated is whether the colour defectives who passed the test will still perform within normal limits when they are fatigued or when carrying out multiple tasks.

It would also be important to follow up with the dispatch centre to make sure that the luminance and chromaticity coordinates of the dispatch test still fall within the range of luminance and chromaticity coordinates of monitors in current use.

There are number of clinical colour vision tests that could also be investigated as predictors for the dispatch test. Three possible tests are (a) Farnsworth-Munsell 100 Hue test (FM 100) (b) Lanthony D15 test and (c) 4th edition HRR pseudoisochromatic plates. The FM 100 Hue test gives a comprehensive assessment of colour discrimination and a finer scale of ones ability to discriminate colous. This may provide a better prediction of who is likely to pass the Dispatch test. The Adams D15 test used in this study was marginally more sensitivity than the D-15 in failing colour-defectives, but not as sensitive Lanthony D15 test. Including a more sensitive test in the test battery may also help in predicting who will pass the Dispatch test. The 4th edition of HRR pseudoisochromatic plates is purported to provide a better grading of the severity of the colour vision defect. It would be interesting to determine whether this improvement in design improves the usefulness of the test in counselling patients regarding their performance on the Dispatch test.

One of the conclusions from the correlation results between the colour differences (ΔE^*) and errors was that ΔE^* between any two colours should be at least 150 units in order to ensure that colour defectives do not make any mistakes. This statement should be tested directly since it has implications in the design of displays for use by individuals with more severe colour vision deficiencies.

APPENDICES

APPENDIX 2.1

MONITOR SET UP AND TEST PROCEDURE

Set-up for the Dell Trinitron Monitor (17 inch)

Brightness settings : buttons on the lower left of the monitor.
Set Brightness to 50
Contrast settings : buttons on the lower right of the monitor
Set Contrast to 80

The rest of the settings have to be made through the menu. This is the centre button on the monitor.
The following settings are actually symbolic representations on the menu and so the descending order of the list corresponds to the descending order on the menu

Size/Centration
- Horizontal Position 36
- Vertical Position 54
- Horizontal Size 67
- Vertical Size 46

Geometry
- Tilt 32
- Pincushion 47
- Barrel 50
- Trapezoid 55
- Parallelogram 51

Colour
- 9300 K

Horizontal Convergence 53
Vertical Convergence 52

Option
- Degauss ON
- Moiré 0
- Position HOSD 54
- Position VOSD 0
- Menu Off

To set the graphic card parameters, place the arrow icon anywhere on the background and click the right button. A menu for the ATI Graphics Card (ATI XPERT 98) settings should appear. Click on the Settings Tab. The settings should be

32 Bit Colour Palette
1024 by 768 Pixels
Small Fonts

All the rest of the settings are the default settings

With these setting, the dimensions of the test objects should be
Size of triangles (base by height) 4 mm by 2 mm
Rectangles 10 mm by 3 mm

Test procedure.

1. The CP-Dispatch program (March 20, 2002 version) should be loaded on the computer and a short-cut should be displayed on the desktop. If not, copy the program from the accompanying disc into the C: drive and create a short cut on the desktop.

2. Double click on the icon to enter the program.

3. Enter the employee's name and CP identification. Note the identification must contain 3 initials followed by a four digit number in order for you to begin the program.

4. The first section will be the signal icon colour test. Allow the employee to read the instructions and review instructions with the person to make sure that they understand the test. **Remind them to double check their responses, because once they click on the "Proceed to the next section" button, they cannot go back and change any responses. Also note that if the person clicks on the "Quit" button at any time, the program will end and all results will be lost.**

5. After completing the signal icon colour section, the instructions for the second part which evaluates one's ability to identify the colours used in the track grid display. Allow the employee to read the instructions and review instructions with the person to make sure that they understand the test. You can inform them that there will be 3 parts to this section of the test. **Remind them to double check their responses, because once they click on the "Proceed to the next section" button, they cannot go back and change any responses.**

6. Instructions for the last part of the test will appear after the third screen. This last part evaluates one's ability to identify colours used to code blocking authorities. Again,

allow the employee to read the instructions and review instructions with the person to make sure that they understand the test. You can inform them that there will be 3 parts to this section of the test.
Remind them to double check their responses, because once they click on the "Proceed to the next section" button, they cannot go back and change any responses.
Also, you can point out to them that the orange and red colours are very similar and so they should be careful when making these responses.

7. At the end of this part of the test, a summary of the results will appear. Click on the "Print to Notepad or "Print to WordPad" in order to get a print out. The employee's results will be saved as a text file on the C: drive under their CP identification number.

8. **Scoring.** Go to the authorities results section on the print out. Count the number of times that orange was identified as red. **Subtract this number from the total error score**.

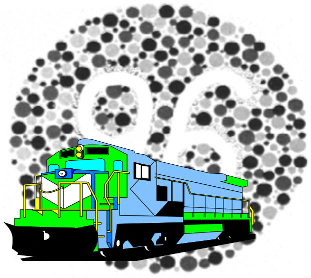
   If this is the first time that the person performed the test or it has been at least 3 months since the last test, then the person is allowed two errors. If it is the second time that the person performed the test within a 3 month period, then only one error is allowed.

   If the person had between 3 and 7 errors and the errors were only those listed in the Table below, then you could repeat the test within a two week period. They are allowed only one mistake on the retest, ignoring any errors where orange is misnamed as red.

Appendix 4.1

Templates of the advertisement used for the study.

COLOUR VISION STUDY



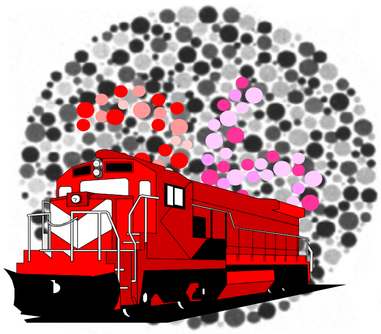
Dr Jeff Hovis from the School of Optometry, University of Waterloo is evaluating colour vision tests designed for the railroad industry. The tests determine one's ability to identify colour codes used to monitor and control train movement.

Individuals with COLOUR VISION PROBLEMS are needed to validate these tests.

The experiment requires between 1 to 2 hours to complete. You will be compensated $10.00 for your time. If you are interested in participating or would like more information, please contact Jeff Hovis at 885-1211 Ext. 6768 or by Email at jhovis@uwaterloo.ca or R. Shankaran at rshankar@sciborg.uwaterloo.ca

This project has received ethics clearance from the Office of Research Ethics at the University of Waterloo (ORE # 9703).

COLOUR VISION STUDY



Dr Jeff Hovis from the School of Optometry, University of Waterloo is evaluating colour vision tests designed for the railroad industry. The tests determine one's ability to identify colour codes used to monitor train movement.

Individuals with COLOUR VISION PROBLEMS are needed to validate these tests.

The experiment requires between 1 to 1.5 hours to complete. You will be compensated $10.00 for your time. If you are interested in participating or would like more information, please contact Jeff Hovis at 885-1211 Ext. 6768 or by Email at jhovis@uwaterloo.ca.

This project has received ethics clearance from the Office of Research Ethics at the University of Waterloo (ORE # 9703).

APPENDIX 5.1

Repeatability results for all the criteria listed in Table 5.5.

The results are for both the colour-normals and colour-defective groups who repeated the test.

Appendix 5.1. Repeatability results for all the criteria listed in Figure 5.5. The results are for both the colour-normals and colour-defective groups who repeated the test.

1. Allow only 1 error on either trial

1/1

Normal
First Trial

Second Trial		Pass	Fail
	Pass	84	7
	Fail	2	0

Proportion Pass both = **0.976**

Proportion Fail both = **0**

1/1

CVD
First Trial

Second Trial		Pass	Fail
	Pass	18	4
	Fail	1	21

K =0.772

Se= 0.095

Proportion Pass 1st but Fail 2nd =**0.053**

Proportion Fail 1st but Pass 2nd =**0.16**

2. Allow only 1 error on either trial, but only with errors listed in Figure 5.5 are permitted

1/1*

Normal
First Trial

Second Trial		Pass	Fail
	Pass	84	7
	Fail	2	0

Proportion Pass both =**0.976**

Proportion Fail both = **0**

1/1*

CVD
First Trial

Second Trial		Pass	Fail
	Pass	15	6
	Fail	2	21

K =0.632

Se=0.117

Proportion Pass 1st but Fail 2nd =**0.118**

Proportion Fail 1st but Pass 2nd =**0.222**

3. Allow 2 errors on either trial

2/2

Normal
First Trial

		Pass	Fail
Second Trial	**Pass**	90	3
	Fail	0	0

Proportion Pass both **=1**

Proportion Fail both **=0**

2/2

CVD
First Trial

		Pass	Fail
Second Trial	**Pass**	20	5
	Fail	1	18

K =0.728

Se= 0.102

Proportion Pass 1st but Fail 2nd **=0.048**

Proportion Fail 1st but Pass 2nd **=0.217**

4. Allow 2 errors on either trial, but only with errors listed in Figure 5.5 are permitted

2/2*

Normal
First Trial

		Pass	Fail
Second Trial	**Pass**	89	4
	Fail	0	0

Proportion Pass both **=1**

Proportion Fail both **=0**

2/2*

CVD
First Trial

		Pass	Fail
Second Trial	**Pass**	16	8
	Fail	2	18

K =0.552

Se= 0.125

Proportion Pass 1st but Fail 2nd **=0.111**

Proportion Fail 1st but Pass 2nd **=0.308**

5. Allow 3 errors on either trial

3/3

Normal
First Trial

Second Trial	Pass	Fail
Pass	92	1
Fail	0	0

Proportion Pass both **=1**

Proportion Fail both **=0**

3/3

CVD
First Trial

Second Trial	Pass	Fail
Pass	22	4
Fail	1	17

K =0.77

Se= 0.096

Proportion Pass 1st but Fail 2nd **=0.043**

Proportion Fail 1st but Pass 2nd **=0.190**

6. Allow 3 errors on either trial, but only with errors listed in Figure 5.5 are permitted

3/3*

Normal
First Trial

Second Trial	Pass	Fail
Pass	91	2
Fail	0	0

Proportion Pass both **=1**

Proportion Fail both **=0**

3/3*

CVD
First Trial

Second Trial	Pass	Fail
Pass	16	6
Fail	1	21

K =0.681

Se=0.109

Proportion Pass 1st but Fail 2nd **=0.058**

Proportion Fail 1st but Pass 2nd **=0.111**

7. Allow one error on first trial and none on second

1/0

Normal
First Trial

Second Trial		Pass	Fail
	Pass	81	6
	Fail	6	0
		Proportion Pass both =0.931	Proportion Fail both =0

K =0.619

Se= 0.120

1/0

CVD
First Trial

Second Trial		Pass	Fail
	Pass	13	6
	Fail	2	23
		Proportion Pass 1st but Fail 2nd =0.133	Proportion Fail 1st but Pass 2nd =0.207

8. Allow one error on first trial and none on second, but only with errors listed in Figure 5.5 are permitted

1/0*

Normal
First Trial

Second Trial		Pass	Fail
	Pass	80	7
	Fail	6	0
		Proportion Pass both =0.93	Proportion Fail both =0

K =0.518

Se= 0.134

1/0*

CVD
First Trial

Second Trial		Pass	Fail
	Pass	11	8
	Fail	2	23
		Proportion Pass 1st but Fail 2nd =0.154	Proportion Fail 1st but Pass 2nd =0.258

9. Allow 2 errors on first trial and 1 on the second

2/1

Normal
First Trial

Second Trial	Pass	Fail
Pass	88	3
Fail	2	0

Proportion Pass both **=0.977**

Proportion Fail both **=0**

K =0.818

Se= 0.086

2/1

CVD
First Trial

Second Trial	Pass	Fail
Pass	19	3
Fail	1	21

Proportion Pass 1st but Fail 2nd **=0.05**

Proportion Fail 1st but Pass 2nd **=0.125**

10. Allow 2 errors on first trial and 1 on the second, but only with errors listed in Figure 5.5 are permitted

2/1*

Normal
First Trial

Second Trial	Pass	Fail
Pass	87	4
Fail	2	0

Proportion Pass both **=0.977**

Proportion Fail both **=0**

K =0.632

Se= 0.117

2/1*

CVD
First Trial

Second Trial	Pass	Fail
Pass	15	6
Fail	2	21

Proportion Pass 1st but Fail 2nd **=0.118**

Proportion Fail 1st but Pass 2nd **=0.222**

11. Allow 2 errors on first trial and 1 on the second-ORANGE-RED not counted as an error

2/1**

		Normal **First Trial**	
		Pass	**Fail**
Second Trial	**Pass**	**91**	**2**
	Fail	**0**	**0**

Proportion Pass both

=1

Proportion Fail both

=0

2/1**

		CVD **First Trial**	
		Pass	**Fail**
Second Trial	**Pass**	**21**	**2**
	Fail	**1**	**20**

K =0.863

Se= 0.075

Proportion Pass 1st but Fail 2nd

=0.045

Proportion Fail 1st but Pass 2nd

=0.090

12. Allow 4 errors on first trial and 1 on the second

4/1

		Normal **First Trial**	
		Pass	**Fail**
Second Trial	**Pass**	**90**	**1**
	Fail	**2**	**0**

Proportion Pass both

=0.978

Proportion Fail both

=0

4/1

		CVD **First Trial**	
		Pass	**Fail**
Second Trial	**Pass**	**21**	**1**
	Fail	**2**	**20**

K =0.863

Se= 0.07

Proportion Pass 1st but Fail 2nd

=0.087

Proportion Fail 1st but Pass 2nd

=0.048

13. Allow 4 errors on first trial and 1 on the second, but only with errors listed in Figure 5.5 are permitted

4/1*

Normal
First Trial

		Pass	Fail
Second Trial	Pass	89	2
	Fail	2	0

Proportion Pass both
=0.978

Proportion Fail both
=0

4/1*

CVD
First Trial

		Pass	Fail
Second Trial	Pass	15	6
	Fail	2	21

K =0.632
Se= 0.117

Proportion Pass 1st but Fail 2nd
=0.118

Proportion Fail 1st but Pass 2nd
=0.222

14. Allow 7 errors on first trial and 2 on the second

7/2

Normal
First Trial

		Pass	Fail
Second Trial	Pass	93	0
	Fail	0	0

Proportion Pass both
=1

Proportion Fail both
=0

7/2

CVD
First Trial

		Pass	Fail
Second Trial	Pass	23	2
	Fail	3	16

K =0.766
Se= 0.098

Proportion Pass 1st but Fail 2nd
=0.115

Proportion Fail 1st but Pass 2nd
=0.111

15. Allow 7 errors on first trial and 2 on the second, but only with errors listed in Figure 5.5 are permitted

7/2*

Normal
First Trial

		Pass	Fail
Second Trial	Pass	91	2
	Fail	0	0

Proportion Pass both =1

Proportion Fail both =0

7/2*

CVD
First Trial

		Pass	Fail
Second Trial	Pass	16	5
	Fail	2	21

K =0.679

Se= 0.111

Proportion pass 1st but Fail 2nd =0.111

Proportion Fail 1st but Pass 2nd =0.192

APPENDIX 6.1

Table A. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by hundred colour-normal subjects. (first trial).

Colour-Normal's Error Distribution in the VDT test **Total Subjects: 100**

VDT 2nd part

Total mistakes: 6

Subject's Response

		Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
VDT colours	**Red**		0	0	0	0	0	0	0
	Yellow	0		0	0	1	0	0	0
	Green	0	0		0	1	0	0	0
	White	0	1	0		0	0	0	0
	Grey	0	0	0	0		0	1	0
	Blue	0	0	0	0	0		0	0
	Purple	0	0	0	0	0	0		0
	Blue Green	0	0	0	0	0	2	0	

VDT 3rd part

Total mistakes: 25

Subject's Response

		Red	Orange	Yellow	Green	Dark Green	Grey	Blue	Light Blue
VDT colours	**Red**		0	0	0	0	0	0	0
	Orange	11		0	0	0	0	0	0
	Yellow	0	1		0	0	0	0	0
	Green	0	0	0		0	0	0	3
	Dark Green	0	0	0	3		0	0	0
	Grey	0	0	0	1	0		0	0
	Blue	0	0	0	0	0	0		2
	Light Blue	0	0	0	0	0	0	4	

Table B. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by fifty two colour-defective subjects (first trial).

Colour-defective's Error Distribution in the VDT test **Total Subjects: 52**

VDT-2nd part

Total mistakes:179

Sum of 52 Subject's Error Responses

		Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
VDT colours	**Red**		0	0	0	0	0	0	0
	Yellow	0		18	0	0	0	0	0
	Green	0	57		0	0	0	1	0
	White	0	1	0		0	0	0	0
	Grey	2	0	1	1		0	8	16
	Blue	0	0	0	0	0		7	0
	Purple	0	0	0	0	0	21		5
	Blue Green	0	0	5	2	22	4	8	

VDT 3rd part

Total mistakes: 224

Subject's Response

		Red	Orange	Yellow	Green	Dark Green	Grey	Blue	Light Blue
VDT colours	**Red**	0	6	0	0	0	0	0	0
	Orange	64	0	0	0	0	0	0	0
	Yellow	0	1	0	25	0	0	0	0
	Green	0	6	66	0	1	0	0	0
	Dark Green	0	0	0	19	0	18	0	0
	Grey	0	0	0	1	16	0	0	1
	Blue	0	0	0	0	0	0	0	0
	Light Blue	0	0	0	0	0	0	0	0

Table C. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by twenty four colour-defective subjects those who passed the VDT test (first trial). Note here that there were no dichromats.

PASSED COLOUR_DEFECTIVE SUBJECTS(24) ERROR DISTRUBUTION

VDT-2nd part

Total mistakes:1

Sum of 24-Passed Subject's Error Responses

		Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
VDT colours	**Red**		**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Yellow	**0**		**0**	**0**	**0**	**0**	**0**	**0**
	Green	**0**	**0**		**0**	**0**	**0**	**0**	**0**
	White	**0**	**1**	**0**		**0**	**0**	**0**	**0**
	Grey	**0**	**0**	**0**	**0**		**0**	**0**	**0**
	Blue	**0**	**0**	**0**	**0**	**0**		**0**	**0**
	Purple	**0**	**0**	**0**	**0**	**0**	**0**		**0**
	Blue Green	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

VDT 3rd part

Total mistakes: 21

Subject Response

		Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
VDT colours	**Red**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**
	Orange	**13**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Yellow	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**
	Green	**0**	**0**	**2**	**0**	**0**	**0**	**0**	**0**
	Dark green	**0**	**0**	**0**	**2**	**0**	**0**	**0**	**0**
	Grey	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Blue	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Light Blue	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Table D. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by six protanomalous subjects those who passed the VDT test (first trial). Note here that there were no dichromats.

PASSED COLOUR_DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Protanomalous(=protan) errors only

VDT-2nd part

Total mistakes: 0

Sum of 6 Passed Subject's Error Responses

		Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
VDT colours	**Red**		**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Yellow	**0**		**0**	**0**	**0**	**0**	**0**	**0**
	Green	**0**	**0**		**0**	**0**	**0**	**0**	**0**
	White	**0**	**0**	**0**		**0**	**0**	**0**	**0**
	Grey	**0**	**0**	**0**	**0**		**0**	**0**	**0**
	Blue	**0**	**0**	**0**	**0**	**0**		**0**	**0**
	Purple	**0**	**0**	**0**	**0**	**0**	**0**		**0**
	Blue Green	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

VDT 3rd part

Total mistakes: 3

Subject Response

		Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
VDT colours	**Red**		**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Orange	**1**		**0**	**0**	**0**	**0**	**0**	**0**
	Yellow	**0**	**0**		**1**	**0**	**0**	**0**	**0**
	Green	**0**	**0**	**1**		**0**	**0**	**0**	**0**
	Dark green	**0**	**0**	**0**	**0**		**0**	**0**	**0**
	Grey	**0**	**0**	**0**	**0**	**0**		**0**	**0**
	Blue	**0**	**0**	**0**	**0**	**0**	**0**		**0**
	Light Blue	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

Table E. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by eighteen deuteranomalous subjects those who passed the VDT test (first trial).Note here that there were no dichromats.

PASSED COLOUR_DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Deuteranamalous(=deutan) errors only

VDT-2nd part

Total mistakes: 1

Sum of 18 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red		0	0	0	0	0	0	0
Yellow	0		0	0	0	0	0	0
Green	0	0		0	0	0	0	0
White	0	1	0		0	0	0	0
Grey	0	0	0	0		0	0	0
Blue	0	0	0	0	0		0	0
Purple	0	0	0	0	0	0		0
Blue Green	0	0	0	0	0	0	0	

VDT 3rd part

Total mistakes: 18

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	2	0	0	0	0	0	0
Orange	12	0	0	0	0	0	0	0
Yellow	0	0	0	1	0	0	0	0
Green	0	0	1	0	0	0	0	0
Dark green	0	0	0	2	0	0	0	0
Grey	0	0	0	0	0	0	0	0
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	

Table F. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by twenty eight colour-defective subjects those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS(28) ERROR DISTRUBUTION

VDT-2nd part

Total mistakes:178

Sum of 28 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red		0	0	0	0	0	0	0
Yellow	0		18	0	0	0	0	0
Green	0	57		0	0	0	1	0
White	0	0	0		0	0	0	0
Grey	2	0	1	1		0	8	16
Blue	0	0	0	0	0		7	0
Purple	0	0	0	0	0	21		5
Blue Green	0	0	5	2	22	4	8	

VDT 3rd part

Total mistakes: 203

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	4	0	0	0	0	0	0
Orange	51	0	0	0	0	0	0	0
Yellow	0	1	0	23	0	0	0	0
Green	0	6	64	0	1	0	0	0
Dark green	0	0	0	17	0	18	0	0
Grey	0	0	0	1	16	0	0	1
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table G. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by fifteen protans those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Protan errors

VDT-2nd part

Total mistakes: 101

Sum of 15 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red		0	0	0	0	0	0	0
Yellow	0		14	0	0	0	0	0
Green	0	33		0	0	0	0	0
White	0	0	0		0	0	0	0
Grey	1	0	1	1		0	2	16
Blue	0	0	0	0	0		5	0
Purple	0	0	0	0	0	11		0
Blue Green	0	0	0	2	15	0	0	

VDT 3rd part

Total mistakes: 102

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red		1	0	0	0	0	0	0
Orange	16		0	0	0	0	0	0
Yellow	0	0		21	0	0	0	0
Green	0	2	40		1	0	0	0
Dark green	0	0	0	7		5	0	0
Grey	0	0	0	0	9		0	0
Blue	0	0	0	0	0	0		0
Light Blue	0	0	0	0	0	0	0	

Table H. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by thirteen deutans those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Deutan errors

VDT-2nd part

Total mistakes: 77

Sum of 13 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red		0	0	0	0	0	0	0
Yellow	0		4	0	0	0	0	0
Green	0	24		0	0	0	1	0
White	0	0	0		0	0	0	0
Grey	1	0	0	0		0	6	0
Blue	0	0	0	0	0		2	0
Purple	0	0	0	0	0	10		5
Blue Green	0	0	5	0	7	4	8	

VDT 3rd part

Total mistakes: 119

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	3	0	0	0	0	0	0
Orange	35	0	0	0	0	0	0	0
Yellow	0	1	0	2	0	0	0	0
Green	0	4	24	0	0	0	0	0
Dark green	0	0	0	10	0	13	0	0
Grey	0	0	0	1	7	0	0	1
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table I. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by nine dichromats those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Dichromatic errors only

VDT-2nd part

Total mistakes: 99

Sum of 9 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0	0	0	0	0	0	0	0
Yellow	0	0	6	0	0	0	0	0
Green	0	23	0	0	0	0	1	0
White	0	0	0	0	0	0	0	0
Grey	1	0	0	0	0	0	7	10
Blue	0	0	0	0	0	0	5	0
Purple	0	0	0	0	0	16	0	4
Blue Green	0	0	0	0	14	4	8	0

VDT 3rd part

Total mistakes: 72

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	0	0	0	0	0	0	0
Orange	19	0	0	0	0	0	0	0
Yellow	0	0	0	5	0	0	0	0
Green	0	6	20	0	0	0	0	0
Dark green	0	0	0	0	0	11	0	0
Grey	0	0	0	1	9	0	0	1
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table J. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by five protanopes those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Protanope errors only

VDT-2nd part

Total mistakes:39

Sum of 5 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0	0	0	0	0	0	0	0
Yellow	0	0	2	0	0	0	0	0
Green	0	9	0	0	0	0	0	0
White	0	0	0	0	0	0	0	0
Grey	0	0	0	0	0	0	1	10
Blue	0	0	0	0	0	0	3	0
Purple	0	0	0	0	0	6	0	0
Blue Green	0	0	0	0	8	0	0	0

VDT 3rd part

Total mistakes: 27

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	0	0	0	0	0	0	0
Orange	8	0	0	0	0	0	0	0
Yellow	0	0	0	5	0	0	0	0
Green	0	2	10	0	0	0	0	0
Dark green	0	0	0	0	0	0	0	0
Grey	0	0	0	0	2	0	0	0
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table K. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by ten protanomalous those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Protanomalous errors only

VDT-2nd part

Total mistakes: 62

Sum of 10 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0	0	0	0	0	0	0	0
Yellow	0	0	12	0	0	0	0	0
Green	0	24	0	0	0	0	0	0
White	0	0	0	0	0	0	0	0
Grey	1	0	1	1	0	0	1	6
Blue	0	0	0	0	0	0	2	0
Purple	0	0	0	0	0	5	0	0
Blue Green	0	0	0	2	7	0	0	0

VDT 3rd part

Total mistakes: 75

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	1	0	0	0	0	0	0
Orange	8	0	0	0	0	0	0	0
Yellow	0	0	0	16	0	0	0	0
Green	0	0	30	0	1	0	0	0
Dark green	0	0	0	7	0	5	0	0
Grey	0	0	0	0	7	0	0	0
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table L. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by four deuteronopes those who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Deuteranope errors only

VDT-2nd part

Total mistakes: 60

Sum of 4 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0	0	0	0	0	0	0	0
Yellow	0	0	4	0	0	0	0	0
Green	0	14	0	0	0	0	1	0
White	0	0	0	0	0	0	0	0
Grey	1	0	0	0	0	0	6	0
Blue	0	0	0	0	0	0	2	0
Purple	0	0	0	0	0	10	0	4
Blue Green	0	0	0	0	6	4	8	0

VDT 3rd part

Total mistakes: 45

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	0	0	0	0	0	0	0
Orange	11	0	0	0	0	0	0	0
Yellow	0	0	0	0	0	0	0	0
Green	0	4	10	0	0	0	0	0
Dark green	0	0	0	0	0	11	0	0
Grey	0	0	0	1	7	0	0	1
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

Table M. Types of errors made in the Track grid (2nd part) and Occupancy Authority (3rd part) of the test by nine deuteranomalous subjects who failed the VDT test (first trial).

FAILED COLOUR DEFECTIVE SUBJECTS ERROR DISTRUBUTION

Deuteranamalous errors only

VDT-2nd part

Total mistakes:17

Sum of 9 Subject's Error Responses

VDT colours

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0	0	0	0	0	0	0	0
Yellow	0	0	0	0	0	0	0	0
Green	0	10	0	0	0	0	0	0
White	0	0	0	0	0	0	0	0
Grey	0	0	0	0	0	0	0	0
Blue	0	0	0	0	0	0	0	0
Purple	0	0	0	0	0	0	0	1
Blue Green	0	0	5	0	1	0	0	0

VDT 3rd part

Total mistakes: 56

Subject Response

VDT colours

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0	3	0	0	0	0	0	0
Orange	24	0	0	0	0	0	0	0
Yellow	0	1	0	2	0	0	0	0
Green	0	0	14	0	0	0	0	0
Dark green	0	0	0	10	0	2	0	0
Grey	0	0	0	0	0	0	0	0
Blue	0	0	0	0	0	0	0	0
Light Blue	0	0	0	0	0	0	0	0

APPENDIX 6.2

Table A. Mean chromaticity co-ordinates and luminance of the measured colours.

	x(mean)	y(mean)	z=1-(X+y)	Luminance Y
Red	0.618	0.356	0.026	12.00
Orange	0.579	0.386	0.036	37.75
Yellow	0.409	0.519	0.072	49.00
Green	0.293	0.609	0.099	40.18
White	0.288	0.313	0.400	52.45
Dark Green	0.247	0.411	0.343	20.60
Grey	0.298	0.311	0.392	8.29
Blue	0.146	0.068	0.786	6.34
Purple	0.288	0.154	0.558	8.33
Light Blue	0.190	0.205	0.605	29.00
Blue Green	0.228	0.277	0.495	28.83

APPENDIX 7.1

Table A. Colour differences in L*a*b*colour space for second and third part of the test.

VDT-2nd part

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
L	54.95	97.40	90.14	100.00	46.71	41.37	46.83	79.21
a	72.06	-24.33	-88.78	0.01	3.69	80.90	72.43	-15.19
b	75.20	102.21	91.01	0.00	0.50	-107.26	-45.06	-19.20

VDT-2nd part- (Lab) Colour Differences

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	108.73	165.40	113.48	101.61	183.18	120.53	130.82
Yellow	108.73	0.00	65.82	105.10	117.04	241.01	183.32	123.10
Green	165.40	65.82	0.00	127.53	136.49	265.48	215.35	132.98
White	113.48	105.10	127.53	0.00	53.42	146.58	100.51	32.12
Grey	101.61	117.04	136.49	53.42	0.00	132.67	82.46	42.43
Blue	183.18	241.01	265.48	146.58	132.67	0.00	63.01	135.71
Purple	120.53	183.32	215.35	100.51	82.46	63.01	0.00	96.92
Blue Gree	130.82	123.10	132.98	32.12	42.43	135.71	96.92	0.00

VDT-3rd part

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
L	54.95	87.95	97.40	90.14	68.95	46.71	41.37	79.21
a	72.06	79.43	-24.33	-88.78	-48.49	3.69	80.90	1.26
b	75.20	104.65	102.21	91.01	19.40	0.50	-107.26	-52.91

VDT-3rd part- (Lab) Colour Differences

	Red	Orange	Yellow	Green	Dark gree	Grey	Blue	Light Blue
Red	0.00	44.84	108.73	165.40	133.58	101.61	183.18	148.37
Orange	44.84	0.00	104.22	168.78	154.90	135.23	216.97	176.10
Yellow	108.73	104.22	0.00	65.82	90.83	117.04	241.01	158.26
Green	165.40	168.78	65.82	0.00	84.85	136.49	265.48	170.12
Dark gree	133.58	154.90	90.83	84.85	0.00	59.79	183.15	88.37
Grey	101.61	135.23	117.04	136.49	59.79	0.00	132.67	62.56
Blue	183.18	216.97	241.01	265.48	183.15	132.67	0.00	103.57
Light Blue	148.37	176.10	158.26	170.12	88.37	62.56	103.57	0.00

Table B. Colour differences in Normal cone space for the second part of the test.

VDT-2nd part- LMS Colour Differences For Normals and Dichromats

Normals

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
L	9.897	32.400	24.427	34.550	5.514	4.080	6.418	18.288
M	2.546	17.920	16.706	20.660	3.207	4.059	2.857	12.502
S	0.835	5.860	5.537	68.450	10.668	76.270	31.367	52.942

Protanopic Simulation

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
L	5.103	35.918	33.480	36.949	5.573	1.462	3.089	21.296
M	2.546	17.920	16.706	20.660	3.207	4.059	2.857	12.502
S	0.835	5.860	5.537	68.450	10.668	76.270	31.367	52.942

Deuteranopic Simulation

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
L	9.897	32.400	24.427	34.550	5.514	4.080	6.418	18.288
M	4.912	16.184	12.239	19.515	3.103	5.309	4.447	11.066
S	0.835	5.860	5.537	68.450	10.668	76.270	31.367	52.942

Delta Eab*=[(dL*)2 + (da*)2 +(db*)2]1/2

Normals- Delta E colour difference (2nd part)

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	27.71	20.83	74.21	10.79	75.67	30.73	53.71
Yellow	27.71	0.00	8.07	62.69	31.02	77.15	39.40	49.45
Green	20.83	8.07	0.00	63.84	23.80	74.68	34.40	47.99
White	74.21	62.69	63.84	0.00	66.98	35.57	49.83	23.91
Grey	10.79	31.02	23.80	66.98	0.00	65.62	20.72	45.13
Blue	75.67	77.15	74.68	35.57	65.62	0.00	44.98	28.59
Purple	30.73	39.40	34.40	49.83	20.72	44.98	0.00	26.45
Blu Green	53.71	49.45	47.99	23.91	45.13	28.59	26.45	0.00

Protan- Delta E colour difference (2nd part)

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	34.80	32.06	76.90	9.87	75.54	30.60	55.47
Yellow	34.80	0.00	2.74	62.66	34.06	79.60	44.22	49.60
Green	32.06	2.74	0.00	63.13	31.42	78.67	42.22	49.13
White	76.90	62.66	63.13	0.00	68.03	39.95	53.28	23.50
Grey	9.87	34.06	31.42	68.03	0.00	65.74	20.85	46.05
Blue	75.54	79.60	78.67	39.95	65.74	0.00	44.95	31.76
Purple	30.60	44.22	42.22	53.28	20.85	44.95	0.00	29.83
Blu Green	55.47	49.60	49.13	23.50	46.05	31.76	29.83	0.00

Deutan- Delta E colour difference (2nd part)

	Red	Yellow	Green	White	Grey	Blue	Purple	Blue Green
Red	0.00	25.67	16.94	73.44	10.92	75.66	30.73	53.14
Yellow	25.67	0.00	8.90	62.72	30.28	76.67	38.25	49.42
Green	16.94	8.90	0.00	64.14	21.62	73.93	32.44	47.82
White	73.44	62.72	64.14	0.00	66.72	34.52	48.92	24.01
Grey	10.92	30.28	21.62	66.72	0.00	65.65	20.76	44.87
Blue	75.66	76.67	73.93	34.52	65.65	0.00	44.97	27.91
Purple	30.73	38.25	32.44	48.92	20.76	44.97	0.00	25.50
Blu Green	53.14	49.42	47.82	24.01	44.87	27.91	25.50	0.00

Table C. Colour differences in Dichromatic cone space for the third part of the test.

VDT-3rd part- LMS Colour Differences For Normals and Dichromats

Normals

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
L	9.897	29.490	32.400	24.427	12.640	5.514	4.080	26.883
M	2.546	9.670	17.920	16.706	8.770	3.207	4.059	29.000
S	0.835	8.390	5.860	5.537	17.270	10.668	76.270	85.580

Protanopic Simulation

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
L	5.103	19.028	35.918	33.480	16.602	5.573	1.462	52.833
M	2.546	9.670	17.920	16.706	8.770	3.207	4.059	29.000
S	0.835	8.390	5.860	5.537	17.270	10.668	76.270	85.580

Deuteranopic Simulation

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
L	9.897	29.490	32.400	24.427	12.640	5.514	4.080	26.883
M	4.912	14.833	16.184	12.239	6.815	3.103	5.309	16.608
S	0.835	8.390	5.860	5.537	17.270	10.668	76.270	85.580

Delta Eab*=[(dL*)2 + (da*)2 +(db*)2]1/2

Normals- Delta E colour difference (3rd part)

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	22.17	27.71	20.83	17.79	10.79	75.67	90.39
Orange	22.17	0.00	9.11	9.13	19.07	24.94	72.70	79.62
Yellow	27.71	9.11	0.00	8.07	24.58	31.02	77.15	80.68
Green	20.83	9.13	8.07	0.00	18.43	23.80	74.68	81.02
Dark green	17.79	19.07	24.58	18.43	0.00	11.19	59.80	72.65
Grey	10.79	24.94	31.02	23.80	11.19	0.00	65.62	82.06
Blue	75.67	72.70	77.15	74.68	59.80	65.62	0.00	35.05
Light Blue	90.39	79.62	80.68	81.02	72.65	82.06	35.05	0.00

Protan- Delta E colour difference (3rd part)

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	17.37	34.80	32.06	21.00	9.87	75.54	100.80
Orange	17.37	0.00	18.97	16.32	9.25	15.10	70.34	86.46
Yellow	34.80	18.97	0.00	2.74	24.23	34.06	79.60	82.24
Green	32.06	16.32	2.74	0.00	22.03	31.42	78.67	83.26
Dark green	21.00	9.25	24.23	22.03	0.00	14.01	61.09	79.93
Grey	9.87	15.10	34.06	31.42	14.01	0.00	65.74	92.25
Blue	75.54	70.34	79.60	78.67	61.09	65.74	0.00	57.86
Light Blue	100.80	86.46	82.24	83.26	79.93	92.25	57.86	0.00

Deutan- Delta E colour difference (3rd part)

	Red	Orange	Yellow	Green	Dark green	Grey	Blue	Light Blue
Red	0.00	23.22	25.67	16.94	16.77	10.92	75.66	87.22
Orange	23.22	0.00	4.09	6.36	20.67	26.79	73.10	77.25
Yellow	25.67	4.09	0.00	8.90	24.67	30.28	76.67	79.91
Green	16.94	6.36	8.90	0.00	17.49	21.62	73.93	80.20
Dark green	16.77	20.67	24.67	17.49	0.00	10.40	59.64	70.46
Grey	10.92	26.79	30.28	21.62	10.40	0.00	65.65	79.06
Blue	75.66	73.10	76.67	73.93	59.64	65.65	0.00	27.10
Light Blue	87.22	77.25	79.91	80.20	70.46	79.06	27.10	0.00

APPENDIX 7.2

Multiple Regressions Results

1. VDT test Track Status (2nd part) - Dichromats in respective space (L+M), Delta S and Error rate
2. VDT test Track Status (2nd part) - Dichromats in respective space (L+M), Delta S and Error rate

APPENDIX 7.2

Multiple Regressions Results

1. VDT test Track Status (2nd part) - Dichromats in respective space (L+M), Delta S and Error rate. The (L+M) represents the luminance dimension and the Delta S represents the hue dimension.

Table A.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.349[a]	.122	.089	8.00437
2	.299[b]	.090	.073	8.07401

a. Predictors: (Constant), DeltaS2ndpart, RespectiveDichromatspaceDeltaLsumM2ndpart

b. Predictors: (Constant), RespectiveDichromatspaceDeltaLsumM2ndpart

c. Dependent Variable: DichromatsErrorPercent2ndpart

Table B.

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	470.818	2	235.409	3.674	.032[a]
	Residual	3395.709	53	64.070		
	Total	3866.527	55			
2	Regression	346.289	1	346.289	5.312	.025[b]
	Residual	3520.238	54	65.190		
	Total	3866.527	55			

a. Predictors: (Constant), DeltaS2ndpart, RespectiveDichromatspaceDeltaLsumM2ndpart

b. Predictors: (Constant), RespectiveDichromatspaceDeltaLsumM2ndpart

c. Dependent Variable: DichromatsErrorPercent2ndpart

Table C.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9.452	2.628		3.597	.001
	RespectiveDichromatspaceDeltaLsumM2ndpart	-.157	.062	-.328	-2.518	.015
	DeltaS2ndpart	-.062	.045	-.182	-1.394	.169
2	(Constant)	6.804	1.832		3.715	.000
	RespectiveDichromatspaceDeltaLsumM2ndpart	-.143	.062	-.299	-2.305	.025

a. Dependent Variable: DichromatsErrorPercent2ndpart

2. VDT test Track Status (2nd part)- Dichromats in respective space (L+M), Delta S and Error rate

Table D.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.417[a]	.174	.143	5.77519
2	.365[b]	.133	.117	5.86116

a. Predictors: (Constant), DeltaS3rdpart, RespectiveDichromDeltaLSumM3rdpart

b. Predictors: (Constant), DeltaS3rdpart

c. Dependent Variable: DichromatsErrosPercent3rdpart

Table E.

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	372.761	2	186.380	5.588	.006[a]
	Residual	1767.702	53	33.353		
	Total	2140.463	55			
2	Regression	285.389	1	285.389	8.307	.006[b]
	Residual	1855.075	54	34.353		
	Total	2140.463	55			

a. Predictors: (Constant), DeltaS3rdpart, RespectiveDichromDeltaLSumM3rdpart

b. Predictors: (Constant), DeltaS3rdpart

c. Dependent Variable: DichromatsErrosPercent3rdpart

Table F.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.973	1.398		4.272	.000
	RespectiveDichromDeltaLSumM3rdpart	-.068	.042	-.206	-1.619	.111
	DeltaS3rdpart	-.059	.023	-.326	-2.562	.013
2	(Constant)	4.723	1.183		3.993	.000
	DeltaS3rdpart	-.066	.023	-.365	-2.882	.006

a. Dependent Variable: DichromatsErrosPercent3rdpart

APPENDIX 9.1

Copyright Permission

Visual Neuroscience

From: Adam Hirschberg
To: Shankaran Ramaswamy;
Subject: Re: Visual Neuroscience: Journal Permission (Vis. Neurosci., 455-460.2004)
Date: December 9, 2008 11:46:01 AM

Dear Shankaran,

Thank you for your message. Cambridge authors do not need formal permission to reprint or reuse figures and tables from their article in another article or chapter of which they are also the author, on the condition that the material to be reprinted is original to the Cambridge author and not obtained from a third party.

To view our policy, please go to our website: www.cambridge.org/us
then click on the link to "Rights and Permissions" and further to "Permission Requests from Authors".

If the material you want to reprint is from another source, you must instead contact the original source for permission.

Sincerely,
Adam

Adam Hirschberg
Rights and Permissions Associate
Cambridge University Press
32 Avenue of the Americas
New York, NY 10013-2473
tel.: 212-337-5088 (direct)
tel.: 212-924-3900
(general) fax: 212-691-3239 (general)
email: ahirschberg@cambridge.org
web: www.cambridge.org/us

From: "Shankaran Ramaswamy" <rshankar@uwaterloo.ca>
To <ahirschberg@cambridge.org> cc 12/08/2008 06:50 PM
Subject Visual Neuroscience: Journal Permission (Vis. Neurosci., 455-460.2004)

Dear Sir/Madam, Submission: Permission to re-use figures in Thesis I would like to use Figures 2, 3, 4 and 5 of the following article in my PhD thesis. I would like to have your permission for the same. Thanks

Ramaswamy, S. and 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, 455–460.

Regards
Shankaran.

Shankaran Ramaswamy BS Optom, FAAO, PhD Candidate
School of Optometry University of Waterloo
Waterloo, Ontario N2L3G1, Canada

APPENDIX 9.2

Logistic Regression Results

Table A. Classification Table

The classification table compares the predicted values for the dependant variable (VDT test), based on the regression model, with the actual observed values in the data

Table B. Variables in the Equation

Table B presents the effects of the variables (Wald test significance) that are in the regression equation.

Table C. Shows the Summary of the Different steps in the Logistic Model

Table D. Model if Term Removed

Table C & D shows two different methods to estimate the model fitness for the data.

Table E. Correlation Matrix

Gives the correlation between the different tests

Table F. Hosmer and Lemeshow Test

Table G. Contingency Table for Hosmer and Lemeshow Test

Table F & G. basically is used to examine the model calibration for the final model (rather than core model or steps in the modelling process)

Logistic Regression Results

Table A. Classification Table

Classification Table[a]

			Predicted		
			VDT test		
Observed			1(Pass)	2(Fail)	Percentage Correct
Step 1	VDT test	1 (Pass)	22	2	91.7
		2 (Fail)	8	20	71.4
		Overall Percentage			80.8
Step 2	VDT test	1 (Pass)	22	2	91.7
		2 (Fail)	8	20	71.4
		Overall Percentage			80.8
Step 3	VDT test	1 (Pass)	22	2	91.7
		2 (Fail)	8	20	71.4
		Overall Percentage			80.8
Step 4	VDT test	1 (Pass)	22	2	91.7
		2 (Fail)	7	21	75.0
		Overall Percentage			82.7
Step 5	VDT test	1 (Pass)	22	2	91.7
		2 (Fail)	8	20	71.4
		Overall Percentage			80.8
Step 6	VDT test	1 (Pass)	23	1	95.8
		2 (Fail)	8	20	71.4
		Overall Percentage			82.7

a. The cut value is 0.500

The classification table shows the outcomes of using the logistic regression model. The classification table compares the predicted values for the dependant variable (VDT test) based on the regression model with the actual observed values in the data. For each case, the predicted response is “Fail” if that case’s model-predicted probability is greater than the cut-off value specified in the dialogs (in this case, the default of 0.5) and “Pass” if the model-predicted probability is less than 0.5. From step to step, the improvement in classification indicates how well the model performs. A better model should correctly identify a higher percentage of the cases. In this case, two different steps classified 82.7% correctly. The tests included in each steps are listed in Table B.

Logistics Regression Equation is given by

$$\textbf{logit (p)=}\ \frac{1}{1+e^{-B_0} \text{ x } e^{-B_1(\text{D-15})} \text{ x } e^{-B_2(\text{Adams D-15})} \text{ x } e^{-B_3(\text{Nagel Anomaloscope})} \text{ x } e^{-B_4(\text{Cn Lantern})} \text{ x } e^{-B_5(\text{Ishihara})} \text{ x } e^{-B_6(\text{HRR})}}$$

e^B is the odds ratio for the independent variable B_i and it gives the relative amount by which the odds of the outcome increase (Odds.Ratio greater than 1) or decrease (Odds.Ratio less than 1) when the value of the independent variable is increased by 1 units.

Table B. Variables in the Equation.

Variables in the Equation

		B	S.E.	Wald	df	Sig.
Step 1[a]	Nagel Anomaloscope	.057	.044	1.677	1	.195
	D15	.277	.193	2.059	1	.151
	AdamsD15	.053	.154	.116	1	.733
	Ishihara	-.190	.282	.456	1	.500
	CnLantern2.3	.086	.205	.175	1	.676
	CnLantern4.6	-.026	.205	.016	1	.898
	HRR	.245	.321	.585	1	.444
	Constant	-1.890	1.243	2.313	1	.128
Step 2[a]	Nagel Anomaloscope	.059	.044	1.808	1	.179
	D15	.270	.185	2.126	1	.145
	AdamsD15	.049	.152	.103	1	.748
	Ishihara	-.205	.259	.625	1	.429
	CnLantern2.3	.073	.181	.164	1	.686
	HRR	.252	.318	.629	1	.428
	Constant	-1.992	.965	4.264	1	.039
Step 3[a]	Nagel Anomaloscope	.059	.043	1.845	1	.174
	D15	.298	.167	3.170	1	.075
	Ishihara	-.199	.257	.600	1	.439
	CnLantern2.3	.069	.179	.147	1	.701
	HRR	.307	.269	1.307	1	.253
	Constant	-2.038	.955	4.553	1	.033
Step 4[a]	Nagel Anomaloscope	.061	.044	1.949	1	.163
	D15	.310	.161	3.696	1	.055
	Ishihara	-.169	.242	.487	1	.485
	HRR	.323	.266	1.470	1	.225
	Constant	-1.960	.919	4.550	1	.033
Step 5[a]	Nagel Anomaloscope	.057	.041	1.906	1	.167
	D15	.317	.160	3.941	1	.047
	HRR	.251	.242	1.075	1	.300
	Constant	-2.213	.863	6.581	1	.010
Step 6[a]	Nagel Anomaloscope	.063	.042	2.239	1	.135
	D15	.372	.156	5.676	1	.017
	Constant	-1.656	.640	6.695	1	.010

a. Variable(s) entered on step 1: Nagel Anomaloscope, D15, AdamsD15, Ishihara, CnLantern2.3, CnLantern4.6, HRR.

Table B summarizes the effect of each predictor variables value for the different steps. The value in the B column corresponds to the B value in the equation and S.E. is the standard error of the value. The constant variable in the first column indicates the constant B_0 term in the equation. The magnitude of B, along with the standard error indicates the effect of the predictor variable (clinical tests) on the predicted variable (VDT test outcome).

The Wald test is used to test the statistical significance of each coefficient (B) in the model for each independent variable (clinical tests). The column labelled "Sig" is actually the p value (significance) of the Wald test. This index is used in the following manner. At Step 1, CN Lantern at 4.6m had the lowest Wald test value and the highest p value, hence it was removed for Step 2. Similarly in Step 2, Adams D15 had the least Wald test value and was removed from the model. At final Step (step 6), both the D15 and the Nagel Anomaloscope had high Wald values with low p values which were nearly equal indicating that neither of the tests could be removed from the model and the analysis stops.

Table C. Summary of the different steps in the Logistic Model

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	42.114[a]	.435	.581
2	42.130[a]	.435	.581
3	42.232[a]	.433	.579
4	42.379[a]	.432	.577
5	42.868[b]	.426	.570
6	43.961[a]	.414	.554

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

b. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

-2 Log Likelihood: This measure indicates how well the model fits the data. Smaller -2 Log likelihood ratio values mean that the model fits the data better; a perfect model has a -2 log likehood value of zero.

Cox and Snell's R square and Nagelkerke's R square estimates the what percentage of the dependent variable may be accounted for by all included predictor variables. It has a theoretical maximum value of less than 1.

Nagelkerke's R square is a version of the Cox & Snell R-square that adjusts the scale of the statistic to cover the full range from 0 to 1. The model with the largest Nagelkerke's R square statistic is “best” according to this measure. However, in this analysis, all there is only a marginal difference between the full model at step 1 and the 2 clinical tests at step 6 and so this table suggests that the simplest model would as the best description of the data depending on the other sets of analyses.

Table D. Model if Term Removed

Model if Term Removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	Nagel Anomaloscope	-22.334	2.554	1	.110
	D15	-22.318	2.521	1	.112
	AdamsD15	-21.114	.113	1	.736
	Ishihara	-21.289	.464	1	.496
	CnLantern2.3	-21.144	.174	1	.676
	CnLantern4.6	-21.065	.016	1	.898
	HRR	-21.358	.601	1	.438
Step 2	Nagel Anomaloscope	-22.453	2.776	1	.096
	D15	-22.373	2.617	1	.106
	AdamsD15	-21.116	.101	1	.750
	Ishihara	-21.383	.636	1	.425
	CnLantern2.3	-21.148	.165	1	.685
	HRR	-21.387	.644	1	.422
Step 3	Nagel Anomaloscope	-22.552	2.872	1	.090
	D15	-23.277	4.322	1	.038
	Ishihara	-21.420	.608	1	.435
	CnLantern2.3	-21.190	.148	1	.701
	HRR	-21.789	1.347	1	.246
Step 4	Nagel Anomaloscope	-22.726	3.073	1	.080
	D15	-23.863	5.346	1	.021
	Ishihara	-21.434	.488	1	.485
	HRR	-21.950	1.520	1	.218
Step 5	Nagel Anomaloscope	-22.883	2.899	1	.089
	D15	-24.323	5.779	1	.016
	HRR	-21.981	1.094	1	.296
Step 6	Nagel Anomaloscope	-23.782	3.602	1	.058
	D15	-26.629	9.298	1	.002

Table D shows the different steps and the estimation of the model fit. All variables are tested here to see if they should be removed from the model. The variables chosen by the backward stepwise method should all have significant changes in -2 log-likelihood if they contribute to the model. The change in -2 log-likelihood is generally more reliable than the Wald statistic.[8] If the two disagree as to whether a predictor is useful to the model, the change in -2 log-likelihood is preferred. In this example, at Step 1, removing CN Lantern

4.6m produced the smallest change in the -2 log-likelihood values with the highest p value and so it was dropped in Step 2. The Adams D-15 was dropped in Step 2 for similar reasons. At Step 6, the analysis stopped because dropping either the Nagel Anomaloscope or Farnsworth D15 produced a significant change in the fit of the model ($p<0.10$).

Table E shows the correlation matrix of all variables in the regression equation. Higher correlation values between tests help explain why some tests did not contribute significantly to the predictions. For example, CN Lantern 4.6m and Adams D-15 had high correlations with the Farnsworth D15 test and, therefore, provided little to the overall predictions and were the first tests to be removed from the model.

Table E. Correlation Matrix

		Constant	Nagel Anomaloscope	D15	AdamsD15	Ishihara	CnLantern2.3	CnLantern4.6	HRR
Step 1	Constant	1.000	-.418	.189	.216	.070	.166	-.635	-.403
	Nagel Anomaloscope	-.418	1.000	-.111	-.052	-.185	-.204	.214	.037
	D15	.189	-.111	1.000	-.365	.206	-.021	-.276	.002
	AdamsD15	.216	-.052	-.365	1.000	-.009	.163	-.184	-.542
	Ishihara	.070	-.185	.206	-.009	1.000	-.078	-.396	-.277
	CnLantern2.3	.166	-.204	-.021	.163	-.078	1.000	-.477	-.218
	CnLantern4.6	-.635	.214	-.276	-.184	-.396	-.477	1.000	.159
	HRR	-.403	.037	.002	-.542	-.277	-.218	.159	1.000
Step 2	Constant	1.000	-.374	.020	.133	-.255	-.204		-.397
	Nagel Anomaloscope	-.374	1.000	-.056	-.017	-.111	-.117		.004
	D15	.020	-.056	1.000	-.443	.110	-.184		.053
	AdamsD15	.133	-.017	-.443	1.000	-.088	.090		-.532
	Ishihara	-.255	-.111	.110	-.088	1.000	-.331		-.235
	CnLantern2.3	-.204	-.117	-.184	.090	-.331	1.000		-.164
	HRR	-.397	.004	.053	-.532	-.235	-.164		1.000
Step 3	Constant	1.000	-.373	.081		-.246	-.224		-.386
	Nagel Anomaloscope	-.373	1.000	-.070		-.113	-.107		-.018
	D15	.081	-.070	1.000		.079	-.172		-.219
	Ishihara	-.246	-.113	.079		1.000	-.324		-.334
	CnLantern2.3	-.224	-.107	-.172		-.324	1.000		-.138
	HRR	-.386	-.018	-.219		-.334	-.138		1.000
Step 4	Constant	1.000	-.409	.030		-.334			-.427
	Nagel Anomaloscope	-.409	1.000	-.090		-.164			-.039
	D15	.030	-.090	1.000		.034			-.243
	Ishihara	-.334	-.164	.034		1.000			-.412
	HRR	-.427	-.039	-.243		-.412			1.000
Step 5	Constant	1.000	-.489	.052					-.673
	Nagel Anomaloscope	-.489	1.000	-.093					-.105
	D15	.052	-.093	1.000					-.257
	HRR	-.673	-.105	-.257					1.000
Step 6	Constant	1.000	-.773	-.200					
	Nagel Anomaloscope	-.773	1.000	-.074					
	D15	-.200	-.074	1.000					

Table F. Hosmer and Lemeshow Test

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11.586	8	.171
2	7.177	8	.518
3	8.304	8	.404
4	6.432	8	.599
5	5.653	8	.686
6	10.526	8	.230

Hosmer-Lemeshow goodness-of-fit statistics helps to determine whether the model adequately describes the data. It evaluates the goodness-of-fit by creating 10 randomly ordered groups of subjects and then compares the number actually in the each group (observed) to the number predicted by the logistic regression model (predicted). The 10 ordered groups are created based on their estimated probability; those with estimated probability below 0.1 form one group, and so on, up to those with probability 0.9 to 1.0. Each of these categories is further divided into two groups based on the actual, observed outcome variable (pass, failure). The expected frequencies for each of the cells are obtained from the model (Table G). This statistic is the most reliable test of model fit for SPSS binary logistic regression, because it aggregates the observations into groups of "similar" cases. The statistic is then computed based upon these groups. The Hosmer-Lemeshow statistic indicates a poor fit if the significance value is less than 0.05. Here, all the different steps fit the data adequately and step 5 seems to be the best fit.

Table G. Contingency Table for Hosmer and Lemeshow Test

Contingency Table for Hosmer and Lemeshow Test

		VDT test = 1.00 (pass)		VDT test = 2.00(fail)		Total
		Observed	Expected	Observed	Expected	
Step 1	1	5	4.403	0	.597	5
	2	4	4.168	1	.832	5
	3	5	3.896	0	1.104	5
	4	2	3.582	3	1.418	5
	5	5	3.205	0	1.795	5
	6	1	2.703	4	2.297	5
	7	1	1.611	4	3.389	5
	8	1	.336	4	4.664	5
	9	0	.080	5	4.920	5
	10	0	.017	7	6.983	7
Step 2	1	5	4.393	0	.607	5
	2	4	4.149	1	.851	5
	3	5	3.901	0	1.099	5
	4	3	3.612	2	1.388	5
	5	4	3.207	1	1.793	5
	6	1	2.705	4	2.295	5
	7	1	1.595	4	3.405	5
	8	1	.341	4	4.659	5
	9	0	.079	5	4.921	5
	10	0	.016	7	6.984	7
Step 3	1	5	4.385	0	.615	5
	2	4	4.132	1	.868	5
	3	4	3.927	1	1.073	5
	4	3	3.599	2	1.401	5
	5	5	3.168	0	1.832	5
	6	1	2.710	4	2.290	5
	7	1	1.658	4	3.342	5
	8	1	.327	4	4.673	5
	9	0	.079	5	4.921	5
	10	0	.015	7	6.985	7
Step 4	1	6	5.194	0	.806	6
	2	4	4.091	1	.909	5
	3	3	3.871	2	1.129	5
	4	4	3.578	1	1.422	5
	5	4	3.113	1	1.887	5
	6	1	2.550	4	2.450	5
	7	1	1.247	4	3.753	5
	8	1	.300	4	4.700	5
	9	0	.045	5	4.955	5
	10	0	.011	6	5.989	6
Step 5	1	5	4.299	0	.701	5
	2	6	4.901	0	1.099	6
	3	3	3.874	2	1.126	5
	4	3	3.643	2	1.357	5
	5	3	2.991	2	2.009	5
	6	2	2.593	3	2.407	5
	7	1	1.328	4	3.672	5

	8	1	.300	4	4.700	5
	9	0	.055	5	4.945	5
	10	0	.015	6	5.985	6
Step 6	1	7	6.362	1	1.638	8
	2	5	3.911	0	1.089	5
	3	4	3.845	1	1.155	5
	4	2	3.542	3	1.458	5
	5	3	2.996	2	2.004	5
	6	2	2.458	3	2.542	5
	7	0	.725	5	4.275	5
	8	1	.140	4	4.860	5
	9	0	.016	5	4.984	5
	10	0	.006	4	3.994	4