City Research Online

City, University of London Institutional Repository

Citation: Carmona, M. L. R. (2006). Variability of chromatic sensitivity: fundamental studies and clinical applications. (Unpublished Doctoral thesis, City University London)

This is the submitted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/12058/

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

VARIABILITY OF CHROMATIC SENSITIVITY: FUNDAMENTAL STUDIES AND CLINICAL APPLICATIONS

Maria Luisa Rodriguez Carmona

Doctor of Philosophy

City University
Henry Wellcome Laboratories for Vision Sciences
Department of Optometry and Visual Science

July 2006

Contents

List of figures 7

List of tables 13

Acknowledgements 15

Declaration 17

Abstract 19

1 Introduction 21

1.1 Overview 22

1.2 Thesis synopsis 26

1.3 Anatomy and physiology of colour pathways 28
1.3.1 The eye 28
1.3.2 The retina 30
1.3.2.1 Photoreceptors 32
1.3.2.2 Post-receptoral retinal pathways 35
1.3.3 Visual processing beyond the retina 39
1.3.3.1 The lateral geniculate nucleus 40
1.3.3.2 The visual cortex 42

1.4 Colour vision models 46
1.4.1 Cone spectral sensitivities: evidence for trichromacy theory 47
1.4.2 Opponent mechanisms 50

1.5 Colour measurement and specification 53
1.5.1 The CIE standard colorimetric observer 53
1.5.2 Standard illuminants 57
1.5.3 CIE chromaticity diagrams 58
1.5.4 Chromatic discrimination 63

1.6 Individual differences in colour vision 69
1.6.1 Variation in 'normal' colour vision 69
1.6.2 Congenital colour vision defects 72
1.6.2.1 Incidence of protan and deutan defects 74
1.6.2.2 Incidence of tritan defects 75
1.6.2.3 Phenotypes of congenital colour vision deficiencies 76
1.6.2.4 Monochromacy 81
1.6.3 Acquired colour vision defects 82

2 Methods of Assessing Colour Vision 85

2.1 Introduction 86

2.2 Clinical and occupational colour vision tests 88
2.2.1 Test Battery 88
2.2.2 Plate tests 89
2.2.2.1 Ishihara pseudoisochromatic plates 90
2.2.2.2 American Optical - Hardy, Rand and Rittler test 92
2.2.3 Hue discrimination or arrangement tests 93
2.2.3.1 The Farnsworth-Munsell 100 hue test 93
2.2.3.2 The Farnsworth-Munsell D-15 test 94
2.2.3.3 The City University test (2nd Edition) 96
2.2.4 Lantern tests 97
2.2.5 Anomaloscopes 99
2.2.5.1 Nagel anomaloscope 99

2.3 Evaluation of colour vision tests 102
2.3.1 Material and methods 103
2.3.1.1 Subjects 103
2.3.1.2 Administration of table based colour vision tests 104
2.3.1.3 Setting the norm and calibration of the Nagel anomaloscope 106
2.3.1.4 Statistical analysis 112
2.3.2 Results 114
2.3.2.1 Efficiency of principal colour vision tests 114
2.3.2.2 Pass/fail comparison 119
2.3.2.3 Performance comparison against the Nagel anomaloscope 120

2.4 Discussion 122

2.5 Conclusions 126

3 The CAD test: A New Approach to Colour Assessment and Diagnosis of Colour Deficiency 127

3.1 Introduction 128

3.2 Methods 130
3.2.1 Spatio-temporal luminance noise 130
3.2.2 Effectiveness of the luminance contrast technique 132
3.2.3 Equipment 136
3.2.4 Definition of stimulus parameters 140
3.2.4.1 Effect of stimulus size 140
3.2.4.2 Effect of background luminance 142
3.2.5 Staircase method 144
3.2.6 Procedure of testing 145
3.2.7 CRT monitor calibration 148
3.2.8 Subjects 151

3.3 Results 151

3.3.1 Typical congenital patterns of colour deficiency 151
3.3.2 Analysis of minimal colour deficient observers 156
3.3.3 Repeatability in the measurement 157

3.4 Discussion and conclusions 159

4 Towards Establishing Limits of 'Normal' Chromatic Sensitivity in Human Vision 161

4.1 Introduction 162

4.2 Methods and Subjects 163
4.2.1 Subjects 164
4.2.2 Statistical analysis 165

4.3 Results 166
4.3.1 Study of variability within normal trichromats 166
4.3.2 Validation of the CAD test against the Nagel anomaloscope 175
4.3.2.1 CAD test versus Nagel anomaloscope 177
4.3.2 Comparison of the CAD test against other colour vision tests 183
4.3.3.1 Ishihara test 183
4.3.3.2 AO-HRR test 185
4.3.3.3 D15 and CU tests 187
4.3.3.4 Overall agreements with the Nagel anomaloscope 189

4.4 Discussion 192

4.5 Conclusions 195

5 Colour Vision Assessment in Subjects with Unusual Anomaloscope Matches – a Molecular Genetics Insight into Colour Deficiency 197

5.1 Introduction 198

5.2 Methods and Subjects 205

5.3 Results 210

5.4 Discussion 218

5.5 Further investigations 219

5.6 Conclusions 220

6 Chromatic Sensitivity and Macular Pigment Density in Human Vision 223

6.1 Introduction 224

6.2 Methods 226

6.2.1 Macular pigment measurements 226
6.2.2 Measurement of chromatic sensitivity 230
6.2.3 Subjects 231
6.2.4 Data analysis and statistics 235

6.3 Results 235
6.3.1 Measurement of the spatial distribution of MP 235
6.3.2 MP group comparison 237
6.3.3 Chromatic sensitivity versus MPOD 240

6.4 Discussion 244

6.5 Conclusions 247

7 Summary and Conclusions 251

7.1 Concluding remarks 256

Appendices 261

Appendix A. Ishihara pseudoisochromatic plates. 261

Appendix B. AO-HRR plates 263

Appendix C. Mathematical definition of the parameters used in the analysis of colour vision tests. 264

Appendix D. Misreadings 266

Appendix E. Statistical analysis for MPOD levels in the fovea and periphery for all groups. 267

Appendix F. 3D scatter plots of the Nagel anomaloscope matching parameters. 269

References and bibliography 271

List of figures

Figure 1.1: Electromagnetic spectrum..... 23

Figure 1.2: A schematic representation based on the Munsell system..... 24

Figure 1.3: Cross section of a human eye. 28

Figure 1.4: Schematic diagram of a tranverse section through the retina showing the different nerve cells and their interconnections..... 31

Figure 1.5: Main structural features of a typical rod and cone..... 33

Figure 1.6: Cone mosaic of the rod-free inner fovea of an adult retina. 35

Figure 1.7: Types of retinal ganglion cell receptive field 38

Figure 1.8: Pathway of nerve signals from the retina to the striate cortex in the brain..... 40

Figure 1.9: Structural organisation of the striate cortex, showing blobs and orientation columns.. 43

Figure 1.10: Schematic drawing of concentric 'double-opponent' receptive field cells in the primary visual cortex 45

Figure 1.11: Stockman and Sharpe (2000) cone fundamentals..... 49

Figure 1.12: Results of the hue cancellation experiment (Hurvich and Jameson 1955)..... 51

Figure 1.13: Simple opponent-colour model of normal human colour vision. 52

Figure 1.14: Colour matching functions of the CIE 1931 standard colorimetric observer, for the RGB system (A), and the XYZ system (B). 54

Figure 1.15: The spectral luminous efficiency curves established by the CIE as the standard photometric observer: V'(λ) for scotopic vision, and V(λ) for photopic vision.. 56

Figure 1.16: The CIE colour matching functions for the 1964 supplementary standard colorimetric observer superimposed on the 1931 standard colorimetric observer..... 57

Figure 1.17: The CIE chromaticity diagrams for the XYZ systems of (A) the CIE 1931 and (B) the CIE 1964 standard colorimetric observers.. 59

Figure 1.18: The derivation on the 1931 CIE-x,y chromaticity diagram of dominant wavelength, complementary wavelength and excitation purity. Plot of the CIE standard sources A, B, C and D_{65}. 62

Figure 1.19: Wavelength discrimination curve based on data obtained by Wright and Pitt (1934).. 64

Figure 1.20: The results of Wright's measurements of just-noticeable differences in chromaticity plotted in CIE 1931 chromaticity space.. 65

Figure 1.21: (A) An example of the experimental results obtained by MacAdam showing the standard deviations involved in colour matching a fixed chromaticity (x_0, y_0) along different direction in the 1931 CIE–x,y chromaticity space. (B) Illustration of ellipse parameters..... 66

Figure 1.22: MacAdam's ellipses plotted in 1931 CIE-x,y chromaticity space showing chromatic discrimination based on the dispersion of a series of colour matches..... 67

Figure 1.23: The approximately uniform CIE u',v' chromaticity diagram with a selection of the lines (just noticeable differences) of Fig. 1.20..... 68

Figure 1.24: The average optical density spectrum of the lens absorption of a 20 year old, and that of a 70 year old..... 70

Figure 1.25: Absorption spectrum of the macular pigment. 71

Figure 1.26: Schematic of the X-linked recessive inheritance for deutan and protan defects.. 74

Figure 1.27: Schematic of autosomal dominant inheritance (tritan defects)..... 76

Figure 1.28: Average luminous efficiency functions for protanopes, deuteranopes and tritanopes.. 78

Figure 1.29: Dichromatic lines of constant chromaticity or hue for (A) protanopes, (B) deuteranopes, and (C) tritanopes, plotted on the normal trichromat's 1931 CIE-x,y chromaticity diagram..... 80

Figure 2.1: Ishihara pseudochromatic plates showing: (A) 'transformation design', (B) 'vanishing design' and (C) 'protan/deutan classification'.. 91

Figure 2.2: The American - Optical Hardy, Rand, Rittler pseudosichromatic plate test.. 93

Figure 2.3: (A) Farnsworth D15 test showing a typical sequence ordered by a protanomalous observer. (B) Results plotted on the results form showing typical results for a normal, protan, deutan and tritan observers. 95

Figure 2.4: Photograph of the City University test illuminated with the Macbeth easel lamp..... 97

Figure 2.5: Photograph of the Holmes-Wright lantern used in aviation. 98

Figure 2.6: (A) Photograph of the Nagel anomaloscope. (B) Illustration of the Nagel anomaloscope split field.. 100

Figure 2.7: Schematic of the type of colour matches for protan and deutan colour vision deficient observers obtained on a Nagel anomaloscope..... 102

Figure 2.8: Age frequency distribution of subjects tested in this study.. 104

Figure 2.9: Matching ranges and midpoints for 231 subjects measured on the Nagel anomaloscope. Data for deuteranopes, extreme deuteranomalous, simple deuteranomalous, normal, simple protanomalous, extreme protanomalous and protanopes..... 107

Figure 2.10: Distribution of the midpoint values of the red-green mixture range among for all subjects. 108

Figure 2.11: Histograms showing the percentage distribution of matching ranges among the normal (N), deuteranomalous (DA), and protanomalous (PA) subjects. 109

Figure 2.12: Scatter plot of matching midpoints versus red-green discrimination index (RGI) for all observers 110

Figure 2.13: Comparison of the number of errors on the Ishihara plate test versus RGI for protan and deutan colour vision deficient subjects. 120

Figure 2.14: Comparison of the number of errors on the AO-HRR plate test versus RGI for protan and deutan subjects. 121

Figure 3.1: Diagram showing luminance profiles for individual frames of spatial random luminance modulation (RLM) for the stimulus array (15 by 15 checks). 131

Figure 3.2: An example of dynamic luminance contrast noise. 132

Figure 3.3: The effect of RLM amplitude on thresholds for detection of an achromatic stimulus pattern in the fovea 133

Figure 3.4: The effect of LC noise amplitude on thresholds for the detection of chromatic stimuli. Chromatic detection thresholds for the red-green (RG) and yellow-blue (YB) channels are plotted. (A) Data for a normal trichromat. (B) Data for a deuteranopic observer 134

Figure 3.5: Chromatic threshold contours for a deuteranopic observer are shown on the CIE-x,y chromaticity chart for different RLM noise amplitudes. 135

Figure 3.6: Photograph of experiment set-up and response button box. 137

Figure 3.7: Schematic of the stimulus used to measure chromatic thresholds showing a square array of 15 x 15 checks, subtending 3.3°. The test target subtends 0.8° and is defined having 21 coloured checks in a 5 x 5 square array. 137

Figure 3.8: A schematic representation of some of the sequence of frames used in the stimulus 139

Figure 3.9: Foveal RG and YB chromatic thresholds measured for different target sizes for 5 observers. 141

Figure 3.10: Foveal R, G, Y and B chromatic thresholds measured for different background luminance levels plotted on a log scale for 5 observers. 143

Figure 3.11: Schematic of the rapid dichromatism test. 146

Figure 3.12: (A) Relative intensity output for the red (R), green (G) and blue (B) phosphors of the CRT display. (B) Luminance versus gun voltage calibration for the R, G and B phosphors. 150

Figure 3.13: Example of the region x,y-chromaticity space that can be reproduced on a typical CRT monitor plotted in the 1931 CIE-x,y chromaticity diagram. The chromaticities of the three phosphors are also plotted. 150

Figure 3.14: Frequency plot showing the number of subjects against the direction in colour space (in degrees) with the highest RG chromatic thresholds. 152

Figure 3.15: Chromatic threshold contours plotted on the CIE-x,y chromaticity chart for four protan observers ranging in severity of colour vision loss.. . 153

Figure 3.16: Chromatic threshold contours plotted on the CIE-x,y chromaticity chart for four deutan observers ranging in severity of colour vision loss... 155

Figure 3.17: Results for 78 mild colour deficient observers tested on both the CAD test and the Nagel anomaloscope.. 157

Figure 3.18: Scatter diagram showing the correlation between two measurements of chromatic thresholds performed on the CAD test for 33 observers..... 158

Figure 3.19: Plot of the percentage error in a measurement for RG and YB threshold of chromatic sensitivity.. 158

Figure 4.1: Age distribution of subjects from a random population sample that particiapted in this study.. 165

Figure 4.2: A scatter diagram of measured RG versus YB thresholds for 245 observers, including 128 male and 117 female. 167

Figure 4.3: Plot of RG and YB thresholds for the random population sample shown in Fig. 4.2 including colour deficient subjects making a total of 450 observers..... 168

Figure 4.4: Frequency distributions of the following chromatic thresholds: (A) YB, (B) RG and (C) ratio YB/RG, obtained in 237 observers with 'normal' trichromatic vision. (D) shows the two directions in 1931 CIE-x,y space used to compute both the YB and RG chromatic thresholds..... 170

Figure 4.5: Histograms showing the different distributions in RG thresholds classified into male and female. 171

Figure 4.6: Data showing the 97.5 and 2.5% statistical limits that define the "standard" normal CAD test observer..... 173

Figure 4.7: Effect of age on YB and RG chromatic thresholds.. 174

Figure 4.8: Plot of RG and YB thresholds in CAD normal units for the remaining colour vision deficient subjects in Fig. 4.3..... 175

Figure 4.9: Plots of the chromatic thresholds for the subjects in Table 4.2. These four subjects are not within the 'normal' on the Nagel anomaloscope, however satisfy the SNO on the CAD test. 178

Figure 4.10: Plots of the chromatic thresholds in CIE-x,y diagram and summarised results for two subjects misclassified in Table 4.3. The classifications on the Nagel anomaloscope and on the CAD test are in disagreement for these two subjects.. 181

Figure 4.11: Results of colour deficient observers measured on the CAD test and Nagel anomaloscope. (A) 51 protan and (B) 108 deutan observers..... 183

Figure 4.12: Results of colour deficient observers measured on the CAD test and Ishihara pseudoisochromatic 25 plates. (A) 67 protan and (B) 138 deutan observers..... 185

Figure 4.13: Results of colour deficient observers measured on the CAD test and AO-HRR pseudoisochromatic 20-plate test. (A) 66 protan and (B) 129 deutan observers.. 187

Figure 5.1: Arrangement of visual pigments in a cone photoreceptor showing the enfolded membrane of the cone outer segment packed with photopigment molecules.. 199

Figure 5.2: Theoretical spectral sensitivity curves as a result of optical density changes for the L-cone sensitive pigment. 201

Figure 5.3: Schematic of the tandem array of L- and M-cones on the q-arm of the X-chromosome. (A) The LCR, exons and introns. (B) Schematic of unequal intragenic crossover that would produce hybrid genes 202

Figure 5.4: Matching ranges and midpoints obtained on the Nagel anomaloscope and predicted phenotypical results from genetics for 23 subjects..... 214

Figure 5.5: RG and YB chromatic thresholds measured on the CAD test in 'standard normal observer' (SNO) CAD test units for 23 subjects..... 215

Figure 5.6: Correlation of Nagel anomaloscope (RGI) versus RG thresholds on the CAD test for 23 subjects. 215

Figure 5.7: CAD data for subjects 7 and 8. These two subjects have the same genotypes with a peak separation $\Delta\lambda_{max}$ of 2.5 nm.. 216

Figure 5.8: Peak separation between visual pigments ($\Delta\lambda_{max}$) versus RG chromatic thresholds..... 217

Figure 6.1: (A) Screen dumps showing the appearance of the stimulus employed to measure the optical density of the macular pigment (MP) for two locations centred 0° and 1.25° from fixation. (B) Example of macular pigment optical density (MPOD) profile obtained by measuring MPOD at a number of retinal locations up to ±8°. 229

Figure 6.2: A comparison of spatial profiles of MPOD measured using motion photometry and the MAP test in four subjects. 230

Figure 6.3: (A) Colour detection thresholds for both YB and RG discrimination measured with the CAD test showing typical results for one normal trichromat. Measurements were taken at the fovea and 3° in the periphery. (B) Bar plot showing typical results for one subject at the fovea and in the periphery..... 231

Figure 6.4: 2-D spatial profiles for two subjects plotted using MATLAB (Mathworks, Inc., Massachusetts).. 236

Figure 6.5: Estimates of mean MPOD for each group of subjects at the end of phase II of the supplementation study. Measurements were made at (A) the fovea and (B) 2.5° in the periphery. (C) Shows a plot of the actual differences in mean MPOD values for each supplemented group with respect to the placebo group. 238

Figure 6.6: Comparison of MPOD spatial profiles for the P-C and the P groups measured along the horizontal meridian from +8° to –8° at different times during the study. (A) Shows data measured in the P-C group at the end of

phase II and after 4 months with no supplementation. (B) Similar data for the P group. (C) Shows MP profiles for the P-C and P groups measured at the end of phase II of the study. (D) Shows the percentage change in the transmittance of blue light when the MPOD values in the P-C group are compared with the corresponding measurements in the P group..... 240

Figure 6.7: (A, C) Chromatic detection thresholds measured at the fovea and (B, D) 5° in the periphery plotted against the available MPOD values at the end of phase II..... 241

Figure 6.8: Chromatic detection thresholds plotted against MPOD values measured at the end of phase III of the study when full spatial profiles of MP density were available for P-C and P groups. (A, C) Average foveal and (B, D) average peripheral MPOD values..... 243

Figure 6.9: Analysis of cone contrasts and the corresponding colour signals generated for the mean colour detection thresholds shown as a black contour in Fig. 6.3A. (A) The cone contrast signals generated in each class of cone photoreceptor for each direction in the CIE–(x, y) diagram. (B) The corresponding red-green and yellow-blue colour signals generated in the absence of MP. The effect of the attenuation of blue light by a MP filter with a peak optical density of 1 log is shown..... 247

List of tables

Table 1.1: Incidences of colour vision deficiencies in men and women..... 72

Table 2.1: Test battery carried out at City University..... 88

Table 2.2: Number of subjects that carried out each different colour vision test used in the study..... 103

Table 2.3: Contingency tables comparing the Ishihara plate, AO-HRR and the Farnsworth D15 colour vision tests with the Nagel anomaloscope. 116

Table 2.4: Summarised results for several colour vision tests with reference to the Nagel anomaloscope. The values of sensitivity, specificity, PPV and NPV are shown..... 117

Table 2.5: Comparison (pass/fail) of the various colour vision tests for the same 164 subjects that did all five tests..... 119

Table 3.1: Staircase parameters used to measure threshold discrimination parameters. 145

Table 3.2: Directions in CIE-x,y colour space (in degrees) used for testing normal trichromats, deutan and protan observers..... 147

Table 3.3: Summary of results for protan subjects A-D in Fig. 3.15. For all four subjects the RG and YB chromatic sensitivity thresholds on the CAD test have been computed. The results on the Nagel anomaloscope: red-green range and RGI, the Ishihara, AO-HRR, CU and D15 tests are given for each subject..... 154

Table 3.4: Summary of results for deutan subjects E-H in Fig. 3.16. For all four subjects the RG and YB chromatic sensitivity thresholds on the CAD test have been computed. The results on the Nagel anomaloscope: red-green range and RGI, the Ishihara, AO-HRR, CU and D15 tests are given for each subject..... 155

Table 4.1: Contingency table showing number of subjects that pass/fail the normal observer criteria for both the CAD test and the Nagel anomaloscope. 177

Table 4.2: Subjects that failed the normal on the Nagel and passed the SNO on the CAD test. 179

Table 4.3: Contingency table showing number of subjects classified as having deutan/protan/normal/tritan colour deficiency on the Nagel anomaloscope that are also classified as deutan/protan/normal/tritan on the CAD. 180

Table 4.4: Contingency table showing number of subjects that pass/fail the Ishihara plate test and the normal observer criteria on the CAD test..... 184

Table 4.5: Contingency table showing number of subjects that pass/fail the AO-HRR plate test and the normal observer criteria on the CAD test. 186

Table 4.6: Contingency table showing number of subjects that pass/fail the D15 test and the normal observer criteria on the CAD test.. 188

Table 4.7: Contingency table showing number of subjects that pass/fail the City University test and the normal observer criteria on the CAD test.. 188

Table 4.8: Summary table showing efficiency of different colour vision tests including the CAD test in comparison to the Nagel anomaloscope..... 189

Table 4.9: Comparison (pass/fail) of the various colour vision tests, including the CAD test, for the same 164 subjects. 191

Table 5.1: Summary of the clinical test results for 23 subjects selected for further testing. 209

Table 5.2: Summary of molecular genetic analysis for 23 colour vision deficient observers..... 212

Table 6.1: Schematic diagram showing the various subject groups, the time course of supplementation with carotenoids and the various tests of visual performance carried out at key stages during the trial. 233

Acknowledgements

I would like to deeply thank the various people who, during the several years I have dedicated to this work, provided me the support and encouragement I needed. This thesis could not have been written without their help.

Firstly I must thank John Barbur for his supervision and for the many and varied discussions we had. Having started from a physics background, I have gained a lot from working in a department with a wide variety of research interests. I have enjoyed learning about the field of vision science, and this would not have been possible without John's contagious enthusiasm.

Jennie Birch for her support and teaching me all about colour vision. Alister Harlow for the control programs. Theresa Squire and Helen Walkey for helpful discussions and advice on numerous occasions. Astrid Dempfle for the thorough reading of my thesis, help with statistics and, overall, very useful input. Steve Grupetta for help with drawings in Illustrator.

Carolina Gheri, Grace Walker, Brice Thurin, Valentina Arena, for their extra patience in carrying out additional experiments when needed.

I'm grateful to the UK Civil Aviation Association for involving me in studies of colour vision in aviation and allowing me more time to finish.

I'm also grateful to Wolfgang Schalch for giving me the opportunity to participate in the LUXEA-study trial.

I have to thank all the 472 subjects for their time and patience in taking part in this study.

The environment at City has been great. I have to really thank everyone, especially Carolina, Brice, Vale, Vincent, Cristiano, Christophe, Grace and Steve for being there when needed.

Finally, I would like to thank my parents and my sister for their advice, encouragement and patience through many and long telephone calls, and for their help when I broke my collar bone in the final stages of writing up. I have to thank especially Uli for his constant and caring support; he must be as fulfilled and relieved as I am to see this project completed!

Declaration

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to myself. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

This investigation involved a number of related studies with the principal aim of assessing the variability in chromatic sensitivity in "normal" trichromats and colour deficient observers. An important outcome was the development of a new method for accurate and efficient measurement of chromatic sensitivity and the establishment of reliable statistical limits that describe the distribution of red-green (RG) and yellow-blue (YB) chromatic sensitivity in normal trichromats. Chromatic sensitivity was assessed using a computer-based psychophysical procedure that employs spatiotemporal luminance contrast (LC) masking techniques to isolate the use of colour signals. The colour-defined stimuli were buried in dynamic LC noise and moved diagonally across a square. The subject's task is to indicate one of the four possible directions of stimulus motion by pressing the corresponding response button. The Colour Assessment and Diagnosis (CAD) test was optimised for stimulus size and background luminance level to provide an efficient measure RG and YB chromatic sensitivity. Colour detection thresholds are assessed along 16 directions in chromaticity space, selected to yield maximum information on RG and YB chromatic sensitivity loss and to distinguish between deutan and protan deficiencies.

The CAD test was used to measure chromatic discrimination thresholds in 472 observers. The results made it possible to evaluate the screening and diagnostic efficiency of the CAD test in comparison with a number of established clinical colour vision tests, e.g., Nagel anomaloscope and Ishihara pseudoisochromatic plates. The specificity and sensitivity of the CAD test versus the Nagel anomaloscope (223 subjects) was found to be 100% and 97.5%, respectively. The diagnostic agreement with respect to the Nagel anomaloscope was 0.99; only two subjects showed inconsistent classification deficiency. Chromatic thresholds measured in normal trichromats were also examined and the variability amongst normal colour vision was investigated. The data obtained were used to establish the statistical limits for the standard normal observer on the CAD test. A template based on these limits was generated and provides an efficient way of separating accurately colour deficient from normal trichromats. The high specificity found is consistent with the correct detection of minimal colour deficiencies that sometimes go undetected in other colour vision tests.

Molecular genetic analysis was also carried out in a small group of subjects with unusual colour deficiencies in an attempt to understand the genotype-phenotype relationship between predicted and observed colour vision losses. A comparison of genetic results from DNA sequencing and chromatic detection thresholds measured on the CAD test suggests that the separation of 30 nm between peak sensitivity of L and M normal photoreceptor pigments is not an absolute requirement for an observer to exhibit normal colour vision as assessed with conventional colour vision tests.

The macular pigment (MP) optical density was also measured in 23 subjects selected for their higher than normal levels of pigment as a result of their participation in a carotenoid supplementation trial. This pre-receptoral filter absorbs light preferentially in the short wavelength region of the visible spectrum and has been shown to affect colour matches. Contrary to expectations, the findings of this investigation showed no correlation between YB thresholds and MP optical density levels. A model that explains this finding was produced confirming the observed experimental findings. In addition, the model also predicts a small improvement in RG chromatic discrimination sensitivity with increased MP optical density in the eye.

1 Introduction

This thesis describes a study of human colour vision with emphasis on the different methods that are used to assess colour vision and the factors that may affect colour vision. This chapter is an introduction to topics that are relevant to the investigations presented in the subsequent chapters.

1.1 Overview

Today colour is all around us; it was only a few decades ago that photography, films and TV were still in black and white. The extensive use of black and white media makes it obviously clear that colour is not usually required to carry out many visual functions. Yet, perception of colour greatly enriches our visual experience. Beyond esthetical reasons, colour can enhance visual performance through increased conspicuity, segmentation and coding.

The human visual system is capable of distinguishing between many different colours (Judd and Wyszecki 1975). The perception of colour is a three-way interdependent system which involves: a source of light, the physical properties of the object, and the observer. This system does not simply depend on the reflectance of the object, but also on the spectral composition of the light illuminating it. In addition, colour perception depends on the characteristics of the observer which vary from one person to another. Besides, what makes human colour perception so useful is the roughly invariant perception of object when viewed in context, despite variation in the composition of incident light. Colour vision is therefore not simply a record of the physical parameters of the light reflected from the object's surface but rather a sophisticated abstracting process ultimately achieved by the brain. Therefore, colour is a primary dimension of visual processing that provides additional information on the physical properties of the environment.

The perception of colour represents the coding and interpretation of the reflectance properties of a physical stimulus through sight. Psychophysics is the name given to studies which examine the relationship between stimulus (physics) and response (psychology). It is in the field of visual psychophysics where important research in human colour vision has occurred. Data on human colour vision are characterised by a large variation in responses to constant stimuli between individuals; common to all biological processes. The study of human perception relies on the assumption of an average or 'normal' observer from which the majority do not differ that much. There are some people that do depart significantly from the average observer and these abnormalities in colour vision

can be congenital or acquired. Congenital colour vision deficiencies are most common and more prevalent in men (~8%) than in women (~0.4%). Acquired deficiencies may be caused by aging effects, retinal or general diseases, toxic effects of certain drugs, etc.

The wavelength range to which the human visual system responds extends normally from 380 to 780 nm. The main regions of the electromagnetic spectrum are shown in Fig. 1.1. The infra-red and ultra-violet regions are outside the visible region of the spectrum; they can provide radiant energy that tans the skin or warms the body, but they cannot normally be seen as light. Our ability to see colours is a result of evolution and it must have appeared early in the evolution of vision. For certain living species, the capacity to differentiate only light and dark without colour must have been a handicap. Two types of photoreceptor evolved first, one responding to one region and a second responding best to another region of the visible spectrum (Nathans 1999). With this arrangement, the brain compares two signals in order to distinguish colour; it is the relative rather than the absolute differences that are interpreted. The trivariant system of colour vision as known today, evolved 30-40 million years ago, in the transition from New World monkeys to Old World monkeys, suggested to facilitate the ability to pick out fruit amongst green foliage (Mollon 1989; Osorio and Vorobyev 1996). The basis of this system is the presence of three photoreceptors in the eye responsible for vision under daylight conditions that are sensitive to different parts of the visible spectrum.

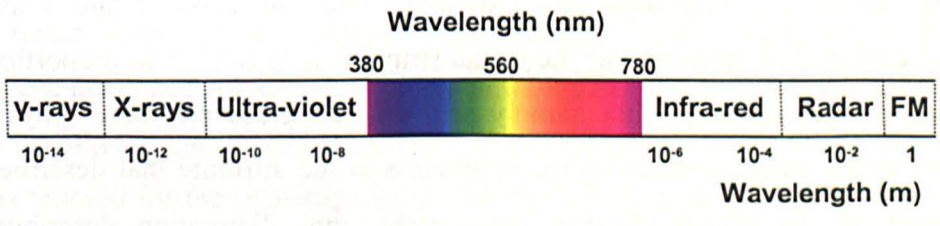


Figure 1.1: Electromagnetic spectrum. Not to scale in order to show the visible spectrum.

In psychophysics there are two main types of experiment; one depends on an observer reporting what is seen or perceived, i.e. colour naming, and the other involves measuring performance thresholds on tasks such as detection, discrimination or colour matching ability. Limitations on colour naming are self-evident as a scientific method since names depend upon a common use of language. Our understanding of the perceptual properties of colour vision has been acquired principally from studies of visual thresholds related to performance-based experiments.

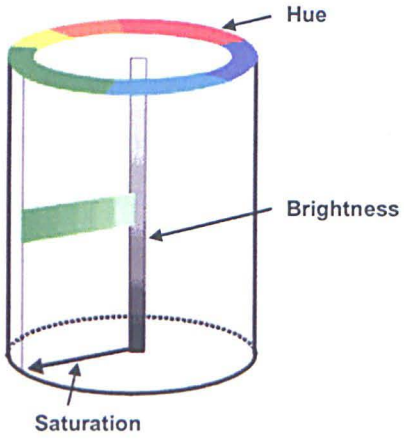


Figure 1.2: A schematic representation based on the Munsell system. The brightness scale is on the vertical axis, the hue scale is on the perimeter of the cylinder and saturation is on a radial scale.

The basic attributes of colour can be described as varying along three perceptual dimensions of hue, brightness and saturation. The dimension of hue is usually described based on how similar the visual stimulus is to one, or to proportions of two of the perceived colours namely red, yellow, green and blue (physically related to the wavelength of light). Brightness is the attribute that describes the lightness of the visual stimulus, i.e. bright, dim. Saturation describes the colourfulness of the colour sample and can be judged with respect to a reference stimulus of the same brightness, i.e. a white stimulus. In 1905, the artist A.H. Munsell designed a colour ordering system with a specific geometrical expansion.

The arrangement of colours was such to produce approximately constant perceptual differences between any two neighbouring samples. Fig. 1.2 shows a simplified representation underlying the basis of a colour ordering system.

Reliable measurements of human perception of colour and in particular the ability to discriminate different colours are becoming increasingly important. Technological improvements in colour displays have made it possible to use a range of colour combinations to improve visual performance for many different tasks. For example, in the transport industry, colour is used for lights, signs, display panel instruments and other control applications. In these situations colour, in addition to enhancing conspicuity (Barbur and Forsyth 1990), is used for the coding, grouping and segmentation of packets of information, which can improve significantly the processing and interpretation of visual information. The conspicuity of objects can be increased significantly by addition of colour, though luminance contrast is undoubtedly the most important parameter. Thus a person with a colour vision deficiency, not able to use correctly colour-coding, may be a handicap in practical situations such as employment. The increased need for clinical testing of congenital colour defective observers has been attributed to the development of demanding occupations and hence the need to screen colour defective observers. In recent years, some employers have considered the need to review their current standards of colour vision. Concern has been expressed particularly by the Civil Aviation Authority in the U.K. Although flight crew must be able to discriminate colours because their work involves the recognition of various colour codes, it is not certain if the present requirements and testing procedures are fully appropriate for the tasks that are carried out.

As a result of technological advances and changes in job complexity in modern transport, many occupational colour vision tests have become less relevant, and in some cases no longer available. There is a need for an objective colour vision test that is tailored for the requirements of specific occupations. It is clear that colour naming and decision making draws in further elements of guessing and subjective judgements rather than certainty of seeing and recognising coloured objects.

1.2 Thesis synopsis

Several investigations that have emanated from the different topics introduced above form the body of this thesis and are described in the remaining chapters.

- This chapter reviews the anatomy and physiology of the human eye with emphasis on aspects relevant to the topics covered in this thesis. A description of the process that starts with the absorption of photons by the eye, converting electromagnetic energy in the form of light into nerve impulses. Following an initial coding within nerve cells in the retina, the nerve impulses are transmitted along the optic nerve and nerve fibres to terminate at the back of the brain. The chapter also describes various methods of colour measurement and specification and introduces the different forms of abnormal colour vision.
- Chapter 2 describes some of the basic clinical techniques available for evaluating congenital colour vision defects. The definitions of the parameters used to evaluate the efficiency of a colour vision test employed throughout this thesis are also included. The results of observers assessed on the principal colour vision tests enables direct comparison with data from a new colour vision test described in Chapter 3.
- Chapter 3 describes the Colour Assessment and Diagnosis (CAD) test that addresses some current problems in colour vision assessment. The process of development and optimisation of the CAD test for measuring chromatic sensitivity is described. This chapter includes also a short investigation into the effects of different stimulus sizes and background luminance on chromatic sensitivity in order to understand the limits of the test. The results obtained in subjects with congenital colour vision deficiency on the CAD test are also described.
- Chapter 4 describes the normal colour vision data that provides a template against which any significant deviation can be assessed. A template defining the standard normal CAD observer is derived from the results of 237 normal trichromats. The CAD test technically provides the basis to investigate the variability within normal colour vision. This chapter also includes a validation of the CAD test against the 'gold standard' test for congenital colour vision deficiency and compares it with the colour vision tests described in Chapter 2.

- In Chapter 5, a study of colour vision assessment in subjects with unusual colour matches on the Nagel anomaloscope is described. In addition to the results on colour matching performance and chromatic discrimination thresholds, genetic sequencing of the photopigment in the photoreceptors has been carried out in the Neitz's lab in Wisconsin (U.S.A). This allows direct comparison of the predicted peak wavelength separation based on genetic identification of their hybrid genes and the measured psychophysical colour discrimination performance.
- Chapter 6 describes an investigation into the effect of the macular pigment, a pre-receptoral filter, on chromatic discrimination thresholds measured in subjects with higher than normal levels of macular pigment.
- Chapter 7 summarises the major findings of all the investigations carried out in this study. The implications of using the results of different colour vision tests in setting minimum colour vision requirements in different occupational environments are also discussed.

1.3 Anatomy and physiology of colour pathways

1.3.1 The eye

The eye is located in the bony orbit close to the midline of the head cushioned by surrounding fatty tissue. Movement of each eye is controlled by three pairs of extraocular muscles that determine the position of the eye within the orbit. A horizontal section of the human eye is shown in Fig. 1.3.

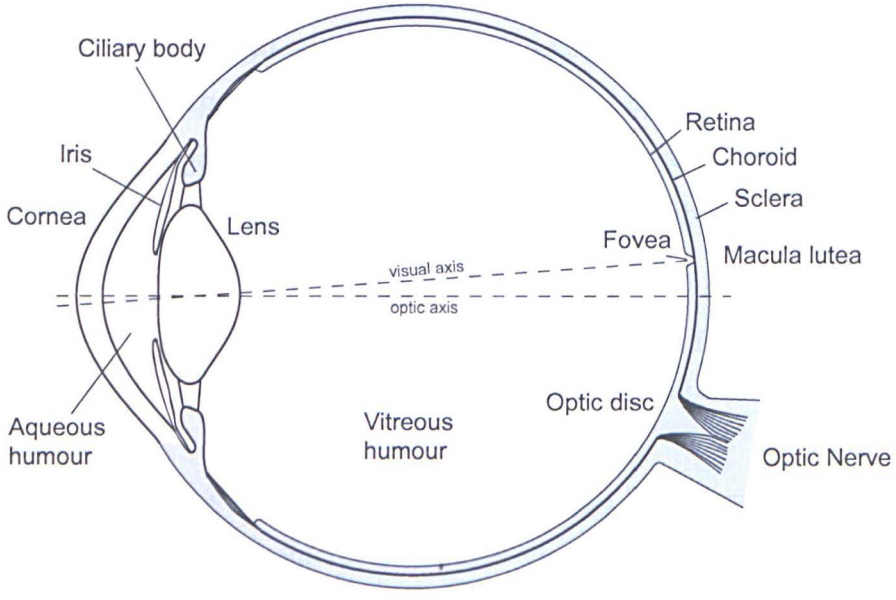


Figure 1.3: Cross section of a human eye.

The outer layer consists of the transparent cornea, the white opaque sclera and a transition region known as the limbal region. The cornea, optically a clear refractive tissue which covers the nearly spherical protuberance at the front of the eye, is the major optical component of the human eye providing clarity, impermeability and protection. Given the absence of blood vessels, protection from microbes and pathogens, provided by the antimicrobial factors in the tear film, is carried out by blinking and the limbal region, also involved in drainage of fluid inside the eye. The iris, the ciliary body and the choroid are located behind the cornea. The iris is a coloured muscular structure whose nearly central aperture

forms the pupil of the eye and determines the amount of light that can pass through. Two involuntary muscles, the sphincter and the dilator control contraction and dilation of the pupil respectively. The pupil diameter can vary from about 2 to 8 mm controlling the light flux captured by the eye. The ciliary body forms a ring of tissue inside the sclera extending from the iris to the anterior edge of the retina. Together, the ciliary muscle and the zonule fibres, change the shape of the lens. The lens of the eye contains crystallins and is elastic in nature whose alteration of its optical properties, during the act of accommodation, allows the viewing of objects at different distances.

The choroid is a vascular layer lying just inside the sclera through which oxygen and nutrients are supplied to the outermost portion of the retina. The retina is a multilayered structure comprising the pigment epithelium, the photoreceptors and other nerve cells, and is responsible for the preliminary processing of light signals. The retina also contains many blood vessels supplying the inner portions of the retina. With the aid of an ophthalmoscope an orange-red hue can be seen that is characteristic of the retina's vascular supply.

The fovea is a small depression from 2.5 to 3 mm in diameter located at the centre of the retina about the visual axis. Apart from having the highest cone density, the foveal depression reduces the amount of scattered light within the retina, which aids the formation of a retinal image of excellent optical quality. The eyeball continuously moves so that light coming from the area on the object of primary interest falls on this region. An image is constantly shifted across different receptor cells by these normal eye movements. The macula lutea (yellow spot) is a central area extending beyond the fovea in which a yellow pigmentation may be seen.

The anterior and posterior chambers located in front of the iris and between the iris and the lens and the ciliary body respectively, are filled with a transparent fluid called the aqueous humour. The aqueous humour circulates through the anterior and posterior chambers. This liquid is continuously generated and absorbed providing a route for nutrient and waste transport via structures (trabecular meshwork) in the limbal region. The area behind the lens is filled with

a transparent gelatinous mass called the vitreous humour. The structure of the eye, which approximates to a sphere with a radius of curvature of about 12 mm, is maintained by the pressure (~10-20mmHg) of these internal fluids.

The visual axis is defined as the line joining the object of interest and the fovea, the area of highest visual resolution (see Fig. 1.3). Slightly more nasally than the visual axis is the optic axis projecting closer to the optic nerve. The optic axis is the longest sagittal distance between the front or vertex of the cornea and the furthest posterior part of the eyeball. It is about the optic axis that the eye is rotated by the extraocular muscles.

Before light can be absorbed in the retina to initiate the visual response it must first transverse the ocular media where a variation in the absorbance with wavelength can change the spectral composition of the stimulus, and thus the resultant colour perception. Structures anterior to the photoreceptors in the retina, acting as pre-receptoral filters, are the cornea, lens and macular pigment (in the macula region) and are described in Section 1.6.

1.3.2 The retina

The retina is contained within the inner limiting membrane, which separates the retina from the vitreous humour, and the choroid, the vascular and pigmented layer attached to the sclera. The retina is less than half a millimetre thick over most of its extent. Within the retina 200 million nerve cells are found that are directly involved with the processing of visual information and are distributed across the retina in a highly organised way. Fig. 1.4 shows the main retinal layers containing the different kinds of nerve cells (neurons) and fibres based on the classification by Polyak (1941). Light passes relatively unblocked through all the retinal layers shown, until the retinal pigment epithelium (RPE), to reach and stimulate the rods and cones.

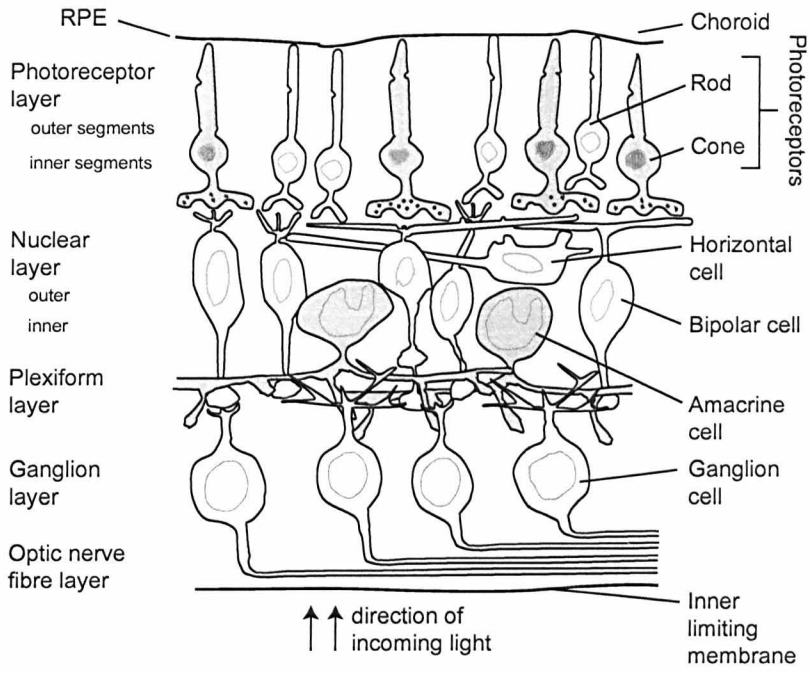


Figure 1.4: Schematic diagram of a tranverse section through the retina showing the different nerve cells and their interconnections. The ultrafine structure of the retina showing the subdivisions of the retina into layers are indicated on the left.

The outermost layer is the pigment epithelium, followed by the photoreceptor layer. The next layer is the outer nuclear layer, which contains nuclei of various types of bipolar cells, horizontal cells and amacrine cells. These cells form intermediate connections in the sequence of signal transportation from the photoreceptors to the ganglion cells. The connections between bipolar cells and ganglion dendritic expansions are found in the plexiform layer. The layer of ganglion cells contains the nuclei of ganglion cells and the axons that constitute the fibres of the optic nerve. Ganglion cell axons traverse the retina and converge at the optic disc where they form the optic nerve to the brain (see Fig. 1.3).

The stream of photons that constitutes the internal visual stimulus enters the retina at the inner limiting membrane and proceeds through the various other retinal layers before it reaches the light sensitive segment of the photoreceptors. The photons that are absorbed elicit signals from the photoreceptors that trigger a chain of events through the neural network of the retina culminating in coded

signals channelled through the axons of ganglion cells to higher levels of visual processing in the brain.

1.3.2.1 Photoreceptors

The pigment epithelium, the outermost layer of the retina, contains cells with processes interdigitating to a small degree the outer segments of photosensitive receptors known as rods and cones and the non photosensitive melanin pigment. The melanin pigment absorbs stray light after it has passed through the neural retina, reducing scatter that would otherwise degrade the retinal image. Fig. 1.5 shows the main structures of the rods and cones. There are roughly 120 million rods and 6 million cones intermingled nonuniformly over most of the retina. A typical rod has a cylindrical outer segment and a cylindrical inner segment, both of similar diameters. A typical cone, on the other hand, is pyramidal and has its largest diameter at the bottom of the inner segment. It is in the outer segment where the interaction of the visual pigment and light initiates the process of vision. Rod outer segments consist of a large number of laminations, referred to as discs. The discs found in the cone outer segments differ from rods in that they are formed by an infolding of the cone membrane. The generation of discs is continuous with discs being replenished in less than two weeks (Young 1978).

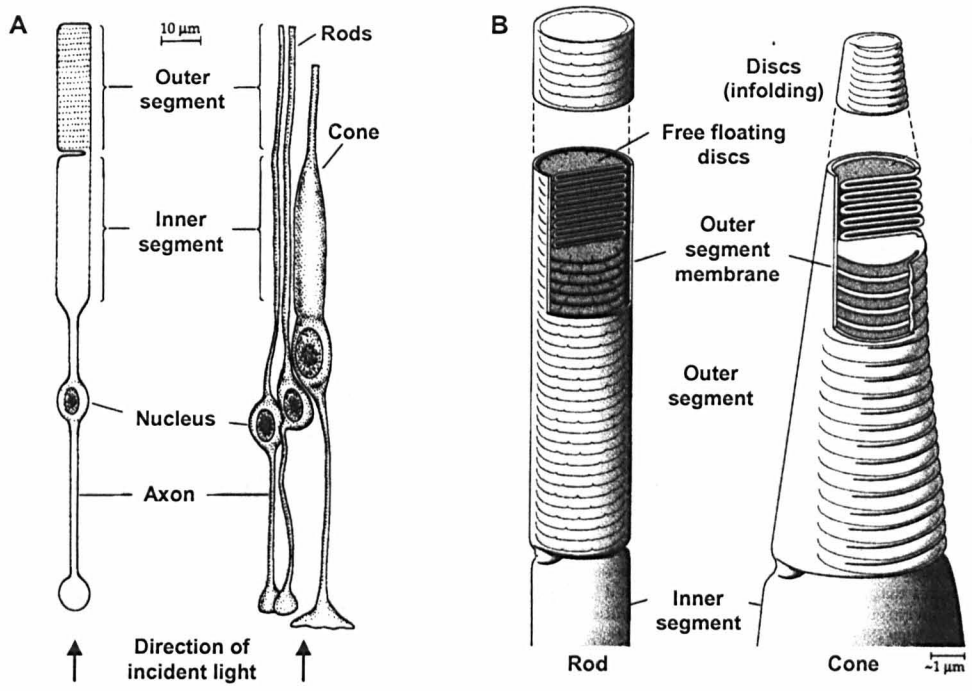


Figure 1.5: Main structural features of a typical rod and cone. (A) Schematic of photoreceptor anatomy, and (B) organisation of discs in rod and cone outer segments. Adapted from Oyster (1999).

The rods mediate vision primarily at scotopic (very low light) levels when visual performance is generally poor and are also involved in motion detection. Cones respond best at photopic (high or daylight) levels, when colour vision and the resolution of fine detail are possible. There are four types of photosensitive pigments in the normal human retina, one contained within the rods and the remaining three distributed amongst the cones. All four photopigments consist of a protein molecule, called opsin, which is bound to the isomer 11-cis retinal (chromophoric group). The chromophoric group is common to all four photopigments, and it is the protein that determines the spectral absorbance properties. The pigment found in rods is known as rhodopsin and it absorbs mostly in the middle wavelength range of the spectrum and transmits light at both the short end and long wavelength ends of the visible spectrum (see Fig. 1.1). When light is absorbed by rhodopsin the bonds in the isomer molecule retinal become rearranged. The molecule changes its shape, untwisting and straightening out in a procedure called phototransduction or *cis-trans* isomerisation, where the

terms *cis* and *trans* refer to two forms of the molecule. This process causes the isomerising molecule to break loose from the protein opsin, by which rhodopsin is bleached and no longer affected by light. The process of isomerisation transmits information to adjacent neural elements within the retina. A similar process is believed to occur during isomerisation of the visual pigments of the cones. The process by which the retinal molecule breaks loose from the opsin gives rise to a change in electrical potential across the outer segment of the rod or cone which increases with the rate at which photons are absorbed.

The time-course of the rod response is longer than that of cones (MacLeod 1972), and the absorption of a single photon by a rod receptor is believed to be sufficient to produce a rod response (Schneeweis and Schnapf 1995). Contrarily, the response of cones is characterised by a fast response and lower photon amplification. There are three types of cone photopigments which are most sensitive or have maximum absorbance values in different parts of the visible spectrum. These are usually referred to as long (L-), medium (M-) and short (S-) wavelength sensitive cones according to the location of the peak of sensitivity in the visible spectrum; traditionally it is common to speak of red, green and blue cones, respectively (see Fig. 1.11). This will be discussed further in Section 1.4.1.

The spatial distribution of cone photoreceptors in the retina affects the quality of colour vision that can be achieved within about ±20° of the visual axis (Hurvich 1981). At larger eccentricities vision becomes virtually monochromatic and serves mainly the detection of novel events. Within the 40° on either side of the visual axis, the ability to see both colour and fine detail increases rapidly as the visual axis is approached. The central region of maximum resolution in the retina, the fovea, subtends a visual angle of about ±2.5°, where the visual axis of the eye intersects the retina. The centre of this region, the foveola, subtends a field of about 1.4° (Wyszecki and Stiles 1982). About 16° from the fovea the optic disc forms the blind spot, an area with no sensitivity to light. In the foveola there are only cones, and outside this the ratio of cones and rods varies continuously until there are nearly all rods and very few cones beyond 40° from the visual axis. In the foveola the cones are thinner and more densely packed than anywhere else in the retina. The relative number of connections from the fovea to the brain is also

greater than from other regions. These features are believed to be responsible for the higher sensitivity of the fovea to visual stimuli as compared to that of other retinal regions, when the eye is light adapted.

The three cone classes are distributed more or less randomly in the retina (Mollon and Bowmaker 1992) and the retinal densities of L-, M- and S-cones can vary significantly from subject to subject. There is also the innermost area of the foveola (~0.34° in diameter) that is free of S-cones (Curcio et al. 1991). Roorda and Williams (1999) combined adaptive optics and retinal densitometry (see Section 1.4.1) to obtain images of the arrangements of S, M and L cones in the living eye. Fig 1.6 is an example of an image of the cone mosaic in the retina. Large individual differences in L:M ratios have been found (such as 1.15 and 3.79) in two observers tested.

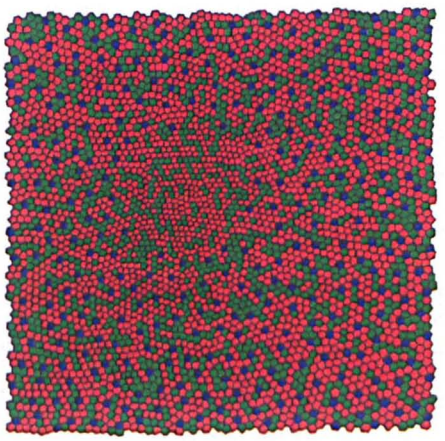


Figure 1.6: Cone mosaic of the rod-free inner fovea of an adult retina. The S-, M-, L-cones have been computer coloured. The diagram shows an inner area free of S-cones and the slightly smaller diameters of the cross sections in the centre compared to the outer edge. Reproduced from Sharpe et al. (1999).

1.3.2.2 Post-receptoral retinal pathways

The photoreceptors synapse onto two major classes of neurones: the horizontal and the bipolar cells. The first light-evoked electrical responses of the retinal

horizontal cells to output from the photoreceptors were recorded by Svaetichin (1953). The responses exhibit graded potential changes (S-potentials) such that the change in potential increases with increasing stimulus intensity. Another feature of the neural coding in the visual system is the duality of polarisation, either an increase of the normal resting potential, hyperpolarisation, or a decrease, depolarisation. These are believed to originate two types of horizontal cells (measured in fish retina, Svaetichin 1958): one type exhibit only hyperpolarisation responses, shows little or no spectral selectivity and is believed to carry information about luminosity, the second type, on the other hand, exhibit both classes of response as the wavelength of the stimulus changes, hence providing the basis for colour vision. More recently, it is thought that horizontal cells provide lateral connections between L and M cones (avoiding S cones) and between S cones and weakly with L and M cones, but do not provide the basis for spectral opponency in the circuitry of the outer retina (Dacey 1999).

Signals from bipolar cell neurones undergo further changes as a result of lateral interaction effects that involve signals from amacrine cells. These are cells that integrate information over larger areas of the retina, up to 0.5° within the macular region and larger in the peripheral retina. Amacrine cells form a network linking inputs from groups of bipolar cells to ganglion cells and other amacrine cells. Several ganglion cells share input from the same L- and M-cones forming complex overlapping synaptic fields. Separate sets of bipolar and amacrine cells are believed to be involved in the extraction of luminance and colour signals. The rod system is also believed to have a separate set of bipolar and amacrine cells, thus achieving further functional separation for responses at scotopic light levels.

Convergence of signals to a ganglion cell occurs over well-defined areas of the retina. These areas are known as receptive fields and they have characteristic spatial distributions across the retina for gathering information about both luminance and spectral modulation. Typical receptive fields are circular in their organisation; where the central receptors provide an excitatory (+) input while the surround receptors form an inhibited (-) annulus. These receptive fields are known as the ON-centre kind. A single receptive field operates locally and cannot provide the coding for brightness, shape and size of a retinal image. This is

achieved by comparison of signals from many adjacent overlapping receptive fields. There are also an equal number of cells with opposite response characteristics, where light falling in the centre inhibits firing of the ganglion cell and illumination of the surround annulus has an excitatory effect. These are known as OFF-centre receptive fields. Their opponent characteristics, adjacent areas of excitatory and inhibitory activity, are fundamental to the way the retina extracts spatial modulations of intensity and spectral content (Lee 2004).

Spatial opponency is an effective algorithm for extracting information about luminance contrast differences. Information about colour is also coded through an opponent mechanism. The existence of spectrally opponent horizontal cells has already been mentioned, but there are also retinal ganglion cells that exhibit such properties. The response to different spectral wavelengths depends on the synaptic connections receiving input from sets of cones containing different photopigments. The spectrally opponent ganglion cells increase their firing rate to one set of wavelengths and decrease it to another, thus making it possible to discriminate coloured lights against a neutral background. These wavelength-opponent cells enhance the boundaries between regions of different colour in a similar way to luminance contrast taking place in other ganglion cells. The separation of luminance and chromatic signals is not complete and some of the receptive fields share and transmit information about luminance and colour at the same retinal location.

There are primarily three classes of retinal ganglion cell: two spectrally antagonistic responses coding for colour and one characterised by a positive response that is independent of wavelength. The antagonistic cells, type I and type II, exhibit an increased firing or positive response to some wavelengths and a decreased firing or inhibitory response to other wavelengths, whereas type III cells give a positive response to a broadband of wavelengths across most of the visible spectrum.

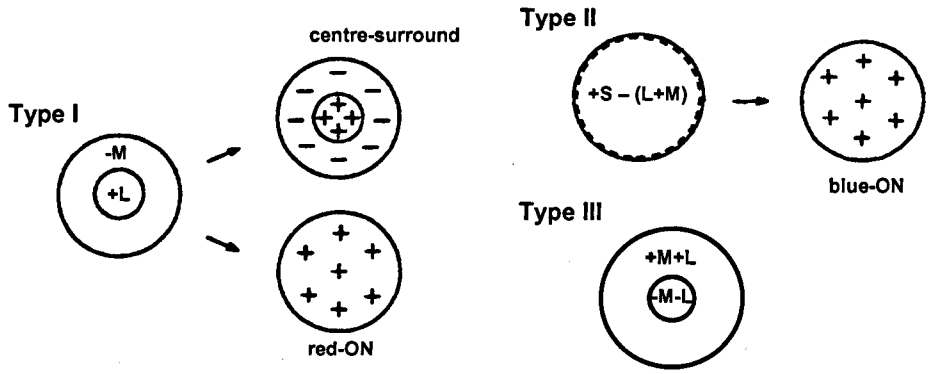


Figure 1.7: Types of retinal ganglion cell receptive field. In type I cells, the receptive field shows a +L/-M centre/surround antagonism and in type II cells, opposing inputs (S-cones vs. L- and M-cones) form two spatially coextensive fields and no centre/surround antagonism. In type III, cells have a concentric broadband receptive field. Adapted from Dacey (1996).

Fig. 1.7 shows the main retinal ganglion cell types based on the way in which the inputs from the three classes of cones are combined. The majority have an antagonistic response to M and L wavelengths (type I), the rest exhibiting an antagonism between S and L+M wavelengths (type II). Type III cells exhibit antagonism between centre and surround but there is no antagonism between cone mechanisms. The centre and surround each combine inputs from both M and L cones. These broadband cells respond to the brightness of the centre versus the brightness of the surround. In type I, the interactions within a receptive field are such that paired receptor types (e.g. L-cones and M-cones) oppose each other in their neural coding. For example, the centre of a receptive field coded for L and M wavelengths may be excited by signals from L-cones, whereas the adjacent surround to the receptive field gives an inhibitory response to signals from M-cones. The output of such a +L/-M antagonistic ganglion cell will depend on the relative amount of stimulation it receives from either the L or M-cones. If signals from the L-cones dominate, thus indicating a greater intensity of light of longer wavelengths in the stimulus, then the ganglion cell will be activated. In counterpart if the signals from the M-cones are higher, the effect is an inhibitory influence on the ganglion cell. These cells also respond to achromatic stimuli because M and L cones in the centre and surround absorb white light to similar extents. For example the cell is excited by small spots of white light but is unresponsive to diffuse white light. Similar arrangements hold for -L/+M

ganglion cells. S versus L+M co-extensive receptive fields respond to input from the S-cones versus the combined signals from L and M-cones. These antagonistic response cells are commonly known as red/green and blue/yellow responses when L/M cells and S/(L+M) cells are involved, respectively.

When a stimulus is composed of wavelengths that activate and inhibit in equal amounts, the centre/surround receptive fields will be mutually antagonistic and the ganglion cell will not fire. These neutral points vary amongst these classes of ganglion cells and can be found in almost any part of the spectrum. The ensuing range of spectral sensitivities amongst these groups of ganglion cells gives a heterogeneous array of opponency throughout the retina (Zrenner 1985).

The types of ganglion cell responses described above may be contributing to luminosity and colour contrast effects and seem to form the basis for further visual analysis at cortical level discussed below. Recent reviews on colour processing in the retina include: Lee (2004), Dacey and Packer (2003) and Lennie (2003).

1.3.3 Visual processing beyond the retina

After 50 milliseconds from the absorption of the photon in the outer segment of the receptor cell, the resultant nerve signals leave the retina and pass along the optic nerve, away from the eye towards the brain. Fibres in the optic nerve segregate into two bundles; those from the left half of the retina continue to the left hemisphere of the brain and those from the right half of the retina are routed to the right hemisphere. This involves some anatomical crossover of certain nerve fibres at the optic chiasm as shown in Fig. 1.8. Behind the optic chiasm, signals from each eye meet in a complex cellular body at the base of the brain known as the dorsal lateral geniculate nucleus (dLGN). It is here where fibres of ganglion cells terminate and form synapses with other fibres leading to a region of the brain called the striate visual cortex, also referred to as visual area 1 (V1).

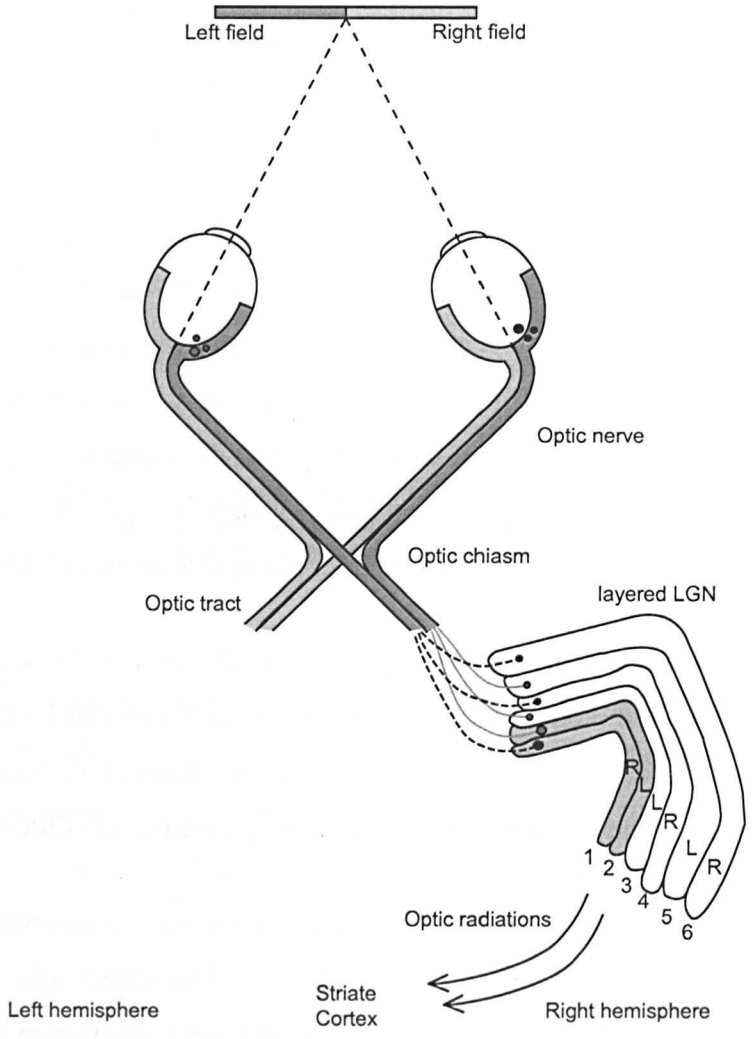


Figure 1.8: Pathway of nerve signals from the retina to the striate cortex in the brain. Nerve fibres from the nasal half of the retina join fibres from the temporal retina of the fellow eye at the optic chiasm. Objects from the left half of the visual field (darker shaded) are processed in the right cortical hemisphere and vice versa. Neurons are layered in the LGN into four parvocellular layers (3-6), two magnocellular layers (1&2) and interlaminar koniocellular layers.

1.3.3.1 The lateral geniculate nucleus

The lateral geniculate nucleus (LGN), shown in Fig. 1.8, has a laminated structure comprising of six distinct layers and several interlaminar regions (Kaas et al. 1978). The receptive field organisation of geniculate neurons in the different layers closely resembles that of retinal ganglion cell inputs to those layers. The

LGN not only receives retinal input but a large number of synapses arise from descending inputs back from the cortex. The LGN therefore acts as a dynamic filter, regulating the passage through to the cortex (Sherman 1996). The top (dorsal) four layers, numbered 3-6, contain neurons with small cell bodies and are referred to as the parvocellular layers. The lower (ventral) two layers, known as the magnocellular layers, contain cells of larger size. The interlaminar regions contain the smallest cell bodies and these are termed the koniocellular layers.

The inputs to the distinct laminar zones of the LGN are segregated according to retinal ganglion cell class. These neurons that project to the LGN are frequently referred to as belonging to the magnocellular or parvocellular pathways. In general, it is the axons of the large (parasol) retinal ganglion cells with fast conduction velocities, which project to the magnocellular layers of the LGN, and the axons of smaller cells (midget), which project to the parvocellular layers. Cells of the magnocellular and parvocellular pathways are often known as M-cells and P-cells, respectively. The broad-band ganglion cells (type III) described above can be either M-cells or P-cells, while the single-opponent cells (type I and II) are exclusively P-cells (Leventhal et al. 1981). Thus, the magnocellular LGN layers are involved only in achromatic vision (Lee et al. 1988). The parvocellular LGN layers relay all colour information to the cortex in addition to information about luminance or achromatic contrast.

There is evidence that neurons in the koniocellular layers form a third pathway that can be traced from the retina to the visual cortex (Casagrande 1994). Small (bistratified) ganglion cells have been found projecting to the parvocellular layers of the primate LGN (Rodieck 1991; Rodieck and Watanabe 1993), but these have also been identified in interlaminar regions of the geniculate layers. These neurons mainly carry signals involving S-cones (blue-ON), suggesting that blue/yellow signals may be part of this third pathway (Dacey and Lee 1994).

Information from the two eyes is segregated into separate layers of the LGN, and there appears to be little or no interocular interaction. The partial decussation of axons at the optic chiasm means that each layer receives only from either the contralateral nasal hemiretina or the ipsilateral temporal retina. Hence signals

initiated in a particular region of the retina are received at the LGN with positional accuracy. Additionally, this segregation might account for stereoscopic vision, where appropriate combination of signals reflect different planes of depth.

In the LGN the coding of signals and the receptive field characteristics are slightly more complex with 'sharper' spatial analysis than those of the retinal ganglion cells, although there is evidence of the same antagonistic centre/surround receptive field (Hubel and Wiesel 1968).

1.3.3.2 The visual cortex

The visual cortex (V1) is the region of the brain that is largely concerned with the processing of visual signals. The nerve fibres terminating in V1, in the occipital lobe of the cerebral cortex, arriving from the LGN are termed optic radiations (see Fig 1.8). Signals coming from objects in different parts of the visual scene remain segregated along the neural pathways as far as V1. The spatial distribution of nerves can be mapped topographically on the surface of the brain corresponding to specific locations on the retina. This topographic mapping is such that the centre region of the visual field has a disproportionably large area of the striate cortex allocated to it. This accounts, at least in part, for the superior visual performance associated with the fovea.

V1 acts as a distributor of signals to other adjacent areas for further processing. Although knowledge of the physiology of colour vision beyond V1 is incomplete, several extra striate cortical areas have been labelled as V2 to V5. Each area appears to be responsive to different basic attributes of a stimulus, i.e. movement of an object, orientation, colour, form, etc. Research into the response characteristics of the cells in these specialised areas of the visual cortex by Zeki (1978; 1980) has led to a theory of functional specialisation of the brain which states that different attributes of the visual scene (such as movement, form, colour) are processed in different regions of the visual cortex. Colour selective cells have been found in areas 1, 2, 3 and 4, although it has been speculated that V4 is primarily responsible for the processing of colour (Zeki 1993). Visual areas

2, 3, and 4 receive input from the parvocellular system. The specific role of these different areas is yet unclear. Area V5, also known as area MT (middle temporal cortex), an area that contains cells sensitive to movement, receives input from the magnocellular system which carries stereoscopic and motion information (Zeki 1974).

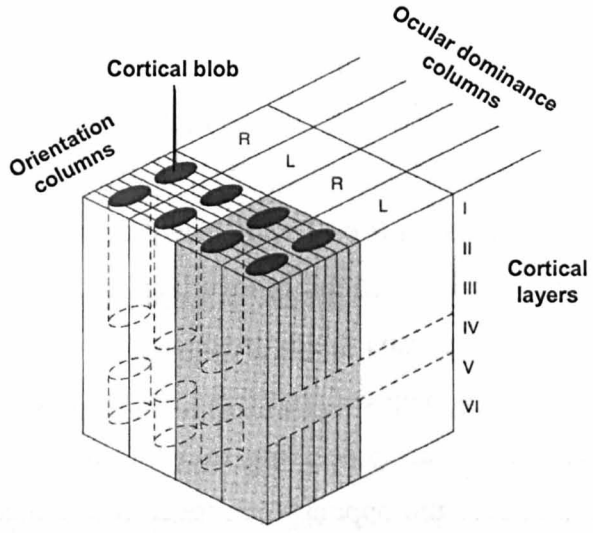


Figure 1.9: Structural organisation of the striate cortex, showing blobs and orientation columns. Reproduced from Schwartz (2004). R and L means corresponds to the right and left eye, respectively.

One characteristic of cells in these areas is that those responsive to the same type of stimulus, e.g. orientation or wavelength, are grouped together in parallel columns perpendicular to the surface of the brain called columns. This specific feature of the organisation of the visual cortex leads to orientation selectivity (Hubel and Wiesel 1974). Another feature of cells in a column is ocular dominance (Hubel and Wiesel 1968). These columns are responsible in distinguishing inputs from the left or right eye. The ocular dominance columns are significantly wider than orientation columns. The inputs from the two eyes remain segregated in the ocular dominance columns of layer IV (see Fig. 1.9). A further organisational element of the visual cortex is the presence of localised zones known as blobs and interblobs. These blobs were identified because the areas are

rich in cytochrome oxidase and traverse the layers of the visual cortex as shown in Fig. 1.9. Chromatically sensitive V1 cells are clustered together in blobs while cells in the interblob regions exhibit orientation specificity (Livingstone and Hubel 1984; Conway 2003; Kiper 2003).

A number of cells, with distinct physiological properties in terms of receptive field organisation that may well be responsible for further processing of colour signals have been identified in cytochrome oxidase rich areas of V1. These cortical neurones have been labelled double antagonistic or double opponent centre/surround receptive fields (Michael 1978). Thus for example (see Fig. 1.10), the central receptive field of one such cell may be activated by long wavelengths and inhibited by medium wavelengths (i.e. +L-M), while its surround would have opposite response characteristics (i.e. -L+M). The properties of these cells appear to be the consequence of interactions of overlapping pairs of different single opponent centre/surround receptive fields from inputs of the LGN through inter-connecting nerve cells. It is believed that these double opponent cells may play important dual roles for colour processing in V1 (Hurlbert 2003). It has been suggested that the inputs to the opponent centre/surround receptive field are not perfectly balanced. A representation of this is shown on the right in Fig. 1.10. A slight asymmetry in the weighting can account for double-opponent cells in responding to both luminance and colour. More importantly, some neurones in V1 may be influenced by light outside their receptive fields as well as by light within their receptive fields. Retinal mechanisms, lateral interactions within V1 and higher cortical areas such as V4 can in principle contribute to colour constancy. Colour constancy accounts for our invariant perception of colours of objects under varied conditions of illumination. This is achieved through a series of processes, some of which are rapid and practically instantaneous with the change of illuminant, and others that are slower or adaptive and involve chromatic adaptation. When all these processes are involved colour constancy can be extremely good and may only be limited by our ability to remember colours (Ling and Hurlbert 2005). Normal functioning of instantaneous colour constancy mechanisms is important since it contributes to image segmentation and object identification and recognition (Barbur et al. 2004).

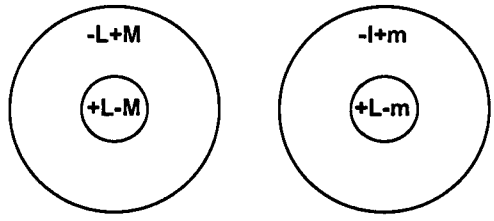


Figure 1.10: Schematic drawing of concentric 'double-opponent' receptive field cells in the primary visual cortex. Centre and surround each contain chromatic opponent mechanisms that generate opposite responses. For example, red-green contrast cells are activated by L-cones in the centre and inhibited by L-cones in the surround. M-cones have the opposite effect; they activate the cell in its surround and inhibit in its centre. Mixed cases of 'L' and 'M' represent imbalance of cone inputs between its centre and surround. Adapted from Hurlbert (2003).

In summary, the efficient coding of signals in the retina transmits information about the local changes in spectral content. The brain analyses the wavelength composition of the light from both adjacent and more distant areas. The special ability to construct constant colours under a wide range of lighting conditions is a consequence of the simultaneous comparison of the light from one region with that of a contiguous region. The long range spatial properties of the receptive fields of cells involved in the colour coding are thus essential in determining how we see colour. Colours are not therefore a property of the real world, but are rather a construct of the brain.

1.4 Colour vision models

Since the mid 18th century the field of colour science has produced an abundance of colour vision theories that aimed to explain how we see colours. Foremost amongst these is a proposition put forward by Newton that colour is a sensation and not a property of the light or electromagnetic radiation. He also realised that the sensation of colour associated with an object is related to the spectral reflectance of an object. Later in 1802 in a lecture at the Royal Society in London, Young added that the human colour vision depended on just three types of colour combinations. The idea he put forward emphasised that these three variables are not a physical property of the light but a physiological limitation of the eye.

This theoretical proposition would become the fundamental basis of what was later known as the trichromatic theory of colour vision. John Dalton around the same time described his own colour vision deficiency. Experimental evidence was provided by Maxwell who demonstrated by means of addition and subtraction of light, that all colours could be matched by the appropriate mixture of the three spectral colours; red, green and blue. Further support came a few years later by Helmholtz who proposed a physiologically based hypothesis of three channels with different but slightly overlapping spectral sensitivities. This experimental evidence showed how certain mixtures of colour were not simply additive in their response to light but appeared to cancel each other. For example, appropriate mixtures of red and green can form a yellow that appears totally devoid of any red or green.

However, the trichromatic theory appeared to contradict the subjective appearance of colours. Hering pointed out that the antagonism between certain colours required a different explanation to that of Young and Helmholtz. He noted that that there were four primary colour sensations: red, green, yellow and blue, but that reddish-green or yellowish-blue could not be perceived together. In 1878 Hering suggested an opponent colour theory, with red opponent to green and yellow opponent to blue. A third opponent channel would process information about lightness and would account for responses of white versus black. These two colour vision theories were thought to be incompatible, but our understanding of

colour vision became enriched when it was suggested that both theories could be accommodated to a zone theory model of colour vision where the mechanisms of trichromacy and opponency occur at different stages of visual processing. Land (1959) put forward the so-called retinex theory in which trichromacy is represented at the receptor level by the three cone types, and opponency holds at some later stage of neural processing.

1.4.1 Cone spectral sensitivities: evidence for trichromacy theory

The theory of trichromacy attributes colour perception to the activity of three primary cone classes. Colour vision is only possible over the operating range of the cones, i.e under photopic or mesopic (twilight) conditions of lighting levels. In scotopic vision humans have no colour discrimination. The basis of trichromacy is the existence of three photoreceptors in the eye with overlapping spectral sensitivities.

A number of different methods have been employed to verify the existence of, and determine the spectral sensitivities of the three types of cone photoreceptors. Direct measurements of photopigment absorption spectra have been carried out using retinal densitometry, microspectroscopy and suction electrode recordings. The method of retinal densitometry involves passing light into the eye and measuring the intensity that is reflected back through the pupil. From such measurements the amount absorbed by the photopigments as a function of wavelength as it travels into and back out of the eye, can be evaluated after accounting for losses due to the absorption by the pigment epithelium. Rushton and collaborators (1955; 1963; 1966) confirmed the existence of the rod photopigment, plus M- and L-wavelength sensitive cone photopigment in the normal eye. To carry out this investigation, people with colour deficiency were used, for example, dichromats (a subject whose retina has one cone photopigment missing, see Section 1.6.2). The technique of retinal densitometry is rather insensitive and because of the fewer number of S- compared to L- and M-wavelength cones it did not yield evidence of S-cone photopigments.

Microspectrophotometry involves measuring the spectral transmission of a small beam going through the outer segment of an individual cone in vitro, compared to the reference beam. Results of measurements using this technique have confirmed the presence of three cone types in the normal human retina plus the rod receptors, containing photopigments with peak sensitivities in different regions of the visible spectrum (Bowmaker and Dartnall 1980). Suction electrode recordings have been obtained from L-cones, which involves drawing the outer segment of a single cone into a micropipette, which is connected to a current-to-voltage converter and then measuring its electrophysiological response when stimulated by light (Kraft et al. 1998).

Indirect methods of measuring cone spectral sensitivities were based on psychophysical data. The spectral sensitivity of the three cone types can be derived from the results of colour matching experiments. Colour matching involves matching one half of a bipartite field of monochromatic test light, with a mixture of altering intensities of three primary lights presented in the other half of the field. For some test wavelengths, a proportion of one or two primaries must be added to the test field to achieve a match. This method of colour matching has revealed the cone sensitivity curves more accurately than by direct methods (Smith and Pokorny 1975). This type of measurement is difficult in colour normals, but more successful in colour deficient observers lacking one or two cone types. In order to compare measurements for colour deficient observers with those of colour normals, the colour deficient observers must have cone sensitivities in their remaining cones that match those in normals. Appropriate observers may be chosen by genetic analysis of the photopigment gene array.

Three classes of cones in the human retina have been isolated from the above techniques. (Dartnall et al. 1983) recorded the absorption spectra transversely across the outer segments of isolated human cones using microspectroscopy. More recent measurements (Smith & Pokorny 1975) using colour matching functions of normals and dichromats have demonstrated conclusively that the three cone pigments, together with that for the rods, each have a maximum absorbance in different parts of the visible spectrum, with a considerable overlap. Ultimately, Stockman and Sharpe (2000) have proposed an improvement to the

cone spectral sensitivities and these are shown in Fig. 1.11. These measurements are consistent with spectral sensitivities measured in dichromats, S-cone monochromats (see Section 1.6.2), and in colour normals, and they reflect typical macular pigment and lens densities for a 2° field.

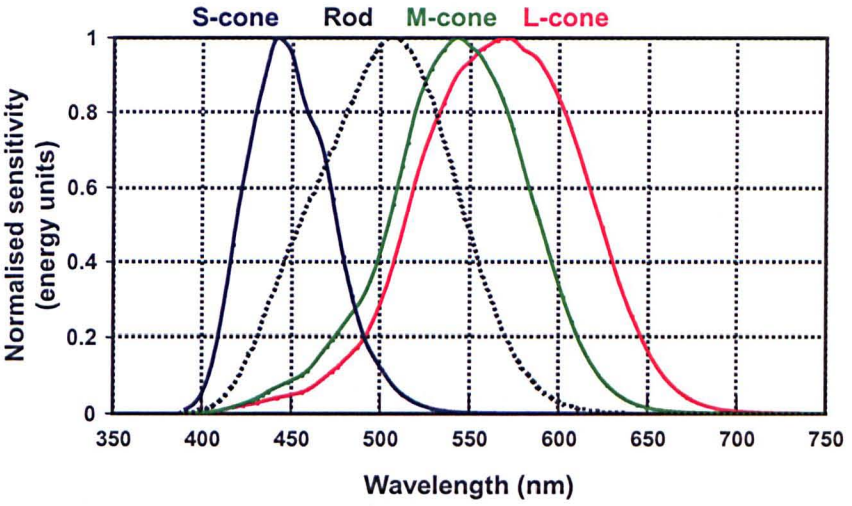


Figure 1.11: Stockman and Sharpe (2000) cone fundamentals. The curves represent normalised spectral sensitivity of the four receptor types in the human retina (Stockman et al. 1999; Stockman and Sharpe 2000). Data obtained from cvrl website (Color vision research laboratories) http://www-cvrl.ucsd.edu/.

The peak sensitivities of the cones lie in the blue (~442 nm), green (~542 nm) and yellow-green (~570 nm) parts of the visible spectrum. The peak sensitivity of the rods occurs at 507 nm. It must be noted that the absorption spectra of the cone pigments represent the probability that an individual photon will be absorbed at the different wavelengths. However, the response of an individual photoreceptor relates only to the number of photons absorbed and signals no information about the wavelength of light; this is known as the principle of univariance (Rushton 1972). Due to the overlapping of the spectra of the three cone pigments, there is a unique combination of absorbance probabilities of each wavelength in the visible spectrum. By comparing the rates of absorption in different classes of cone, the visual system is able to discriminate wavelength.

In summary, detailed evidence of the three cone types has been achieved by microspectroscopy, suction electrophysiology and psychophysical methods. The most important psychophysical experiments have been those of colour matching initiated by Maxwell when he allocates the three colours to the corners of a triangle. In this triangle the corners represent the physiological primaries and any given location in the triangle can in principle specify the relative probabilities that a photon will be absorbed by each of the three types of pigment.

1.4.2 Opponent mechanisms

Experimental evidence supporting the neural opponent-colour theory started to appear in the 1950s. The opponent colour theory, first proposed by Hering, never challenged the initial stages of processing expressed by the trichromatic theory. There is electrophysiological evidence, described earlier, of spectral antagonism from nerve cells responses both in the retina and in the visual pathways of the brain. The psychophysical evidence originates largely from colour mixing studies using hue cancellation technique. In this method pairs of wavelengths that elicit opposite hue responses (i.e. red and green or yellow and blue) are mixed by superposition and their relative energies varied until a ratio is determined at which neither of the opposite hues can be detected in the mixture. The relative energies required for the hue cancellation serve as a measure of the amplitude of the relevant colour-opponent function at each wavelength. For example, the intensities of red and yellow appearing lights were taken to be visually equal when presented in amounts that produced the perception of a balanced orange mixture. Using these hue cancellation paradigms, the psychophysical colour opponent channels were isolated (Jameson and Hurvich 1955). An example of the results of these experiments is shown in Fig. 1.12.

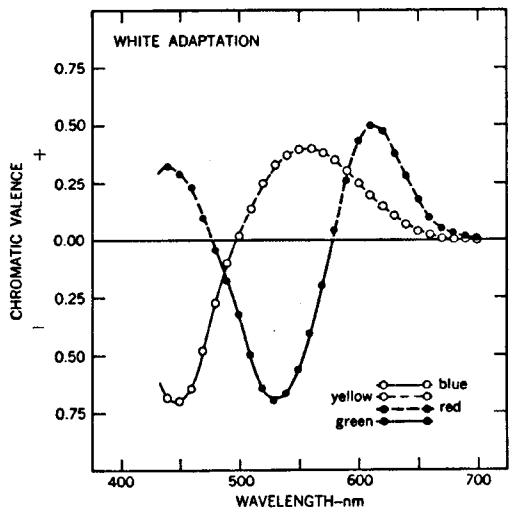


Figure 1.12: Results of the hue cancellation experiment (Hurvich and Jameson 1955). By adjusting the amount of blue or yellow and red or green, any sample wavelength can be matched. Complementary wavelengths can be used to cancel each other for all wavelengths except the four unique hues (blue, green, yellow and red). Chromatic valence on the ordinate, refers to the amount of cancellation component, expressed in relative energy units, required to cancel the opponent hue of test stimuli presented at equal energy at various wavelengths. Red and yellow have been arbitrarily assigned as positive and green and blue as negative, for graphically illustrating the opponent processing. Reproduced from Wyszecki and Stiles (1982).

Several models have emerged that attempt to organize the physiology, anatomy, and psychophysics of colour vision (De Valois and De Valois 1993; Hunt 1991). Fig. 1.13 below shows a simplified model that combines known aspects of physiology and psychology of colour vision. The first stage consists of the S-, M- and L-cone inputs. The second stage is an opponent processing resulting from subtractive processes; the outputs of the L- and M-cones subtract each other to produce the red-green (r-g) chromatic channel and the S-cone signals subtract to those of L+M cones to form the yellow-blue (y-b) chromatic channel. The diagram shows that the S-cones are connected to the r-g channel in order to account for the perception of violet at short wavelengths (Lennie and D'Zmura 1988). The luminance channel, used for the perception of blackness and whiteness, is fed by the sum of the L-cones and M-cones (r+g).

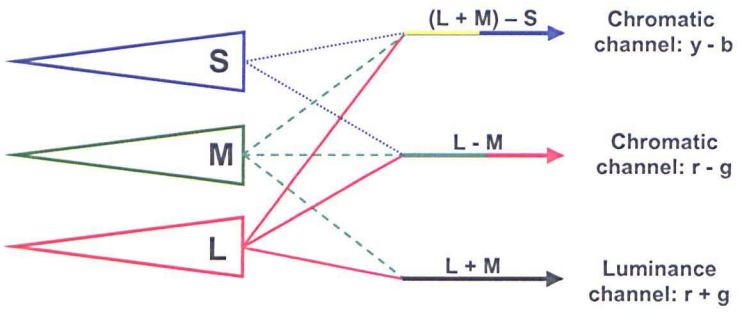


Figure 1.13: Simple opponent-colour model of normal human colour vision.

One of the interesting attributes of this later stage of neural processing in the cortex is the multiplexing of chromatic and achromatic signals. This was originally proposed by Ingling and Martinez-Uriegas (1983) and assumes that the parvocellular pathway is able to carry both chromatic and achromatic information. Which of these is transmitted depends in part on the temporal and spatial characteristics of the stimulus.

1.5 Colour measurement and specification

Colour specification is concerned with the numerical characterisation of colour. In other words, stimuli with the same specification should look the same to an observer, with normal colour vision, in different parts of the world, when viewed under the same conditions. This branch of colour science is known as colorimetry. It is also concerned with the measurement of the smallest colour difference that an observer may perceive when the differences in the spectral radiant distributions of the given visual stimuli are such that a full colour match is not observed.

1.5.1 The CIE standard colorimetric observer

Almost all modern colour measurement is based on the CIE system of colour specification. This is a system set up by the international committee Commission Internationale de l'Eclairage (CIE) in 1931. It is an empirical system based on the fact that observers can match colours with additive mixtures of three reference stimuli, called primaries, in amounts known as tristimulus values. Any set of colour stimuli constitutes a set of primaries if each primary cannot be matched by an additive mixture of the other two, i.e., red, green and blue. This is a consequence of the near linearity of Grassmann's law of additive colour mixtures. Using primaries at specified wavelengths, the CIE defined a standard set of tristimulus values to match each different wavelength of the spectrum.

Suppose **R**, **G**, **B** is a set of primary light sources. If these are used to match a colour stimulus **C** with spectral radiant power distribution $\mathbf{C}_\lambda$, using a mixture of the primaries, so that a match is expressed by: $\mathbf{C}_\lambda = r\mathbf{R}_\lambda + g\mathbf{G}_\lambda + b\mathbf{B}_\lambda$, then the tristimulus values r, g, b are given by:

$$r = k \int_{380}^{780} C_\lambda \bar{r}(\lambda) d\lambda$$

$$g = k \int_{380}^{780} C_\lambda \bar{g}(\lambda) d\lambda$$

$$b = k \int_{380}^{780} C_\lambda \bar{b}(\lambda) d\lambda \qquad (1.1)$$

where $\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)$ are the appropriate colour matching functions (CMF) and k a scaling function, and the limits of λ are 380 to 780 nm (visible region of the electromagnetic spectrum).

The first colour matching experiments were carried out by Wright (1928) and Guild (1931). Due to variations in the properties of the retina from one part to another, the matching size of the angular field was set to a 2° field size. The primary stimuli were monochromatic and of wavelengths, λ_R=700 nm, λ_G=546.1 nm, and λ_B=435.8 nm, chosen to coincide with the two prominent lines in the mercury discharge spectrum facilitating wavelength calibration. The red stimulus was chosen to be in the part of the spectrum where hue changes vary slowly with wavelength. The units of the tristimulus values *r*, *g*, and *b* were defined such that equal proportions of the primaries matched an equal energy white (a stimulus with a flat spectral radiance distribution). In this colorimetric specification, known as the RGB system, the tristimulus values were negative over some parts of the visible spectrum. The CMFs for the RGB system are shown in Fig. 1.14A.

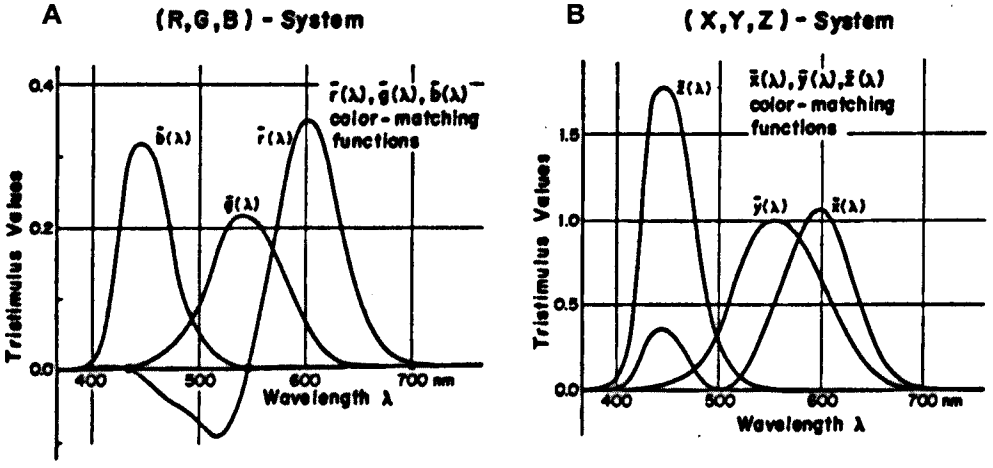


Figure 1.14: Colour matching functions of the CIE 1931 standard colorimetric observer, for the RGB system (A), and the XYZ system (B). Reproduced from Wyszecki and Stiles (1982).

The CIE carried out a further colorimetric transformation of the RGB system to obtain positive tristimulus values over the whole spectrum. This was a linear transformation to ensure the tristimulus values denoted by X, Y, Z and associated CMFs, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$, of any real colour stimulus are never negative. These are shown in Fig. 1.14B.

A further characteristic of the 1931 CIE standard colorimetric system was that the choice of primaries was such that $\bar{y}(\lambda)$ was identical to the relative sensitivity of the eye to light of different wavelengths. The relative sensitivity or luminous efficiency of the eye is known as the V_λ curve, and was adopted as standard by the CIE in 1924.

The sensitivity of the eye varies with wavelength and whether cones (photopic vision) or rods (scotopic vision) are active. At a high luminosity level only the cone function would be found and at low levels only the rod function. Fig. 1.15 below shows the normalised functions equated to unity at the most sensitive wavelength; 555 nm for cones (V_λ) and 507 nm for rods (V'_λ). A change in the relative spectral sensitivity occurs as the illumination level falls, towards a decrease for reddish colours (darker) and an increase for bluish colours. This is known as the Purkinje shift.

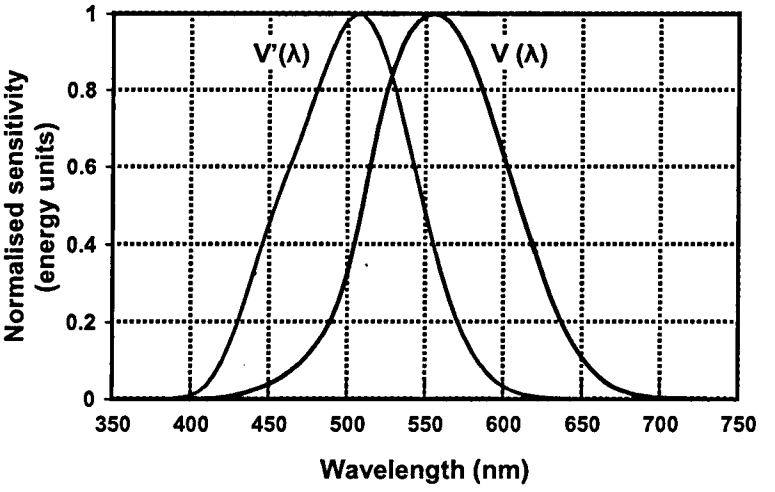


Figure 1.15: The spectral luminous efficiency curves established by the CIE as the standard photometric observer: V'(λ) for scotopic vision, and V(λ) for photopic vision. Data obtained from cvrl website (Color vision research laboratories) http://www-cvrl.ucsd.edu/.

The original 1931 CIE standard observer was based on experiments using a 2° field of view, limiting the influence of rods in colour matching. This is a much narrower field of view than is normally used for critical colour appraisal. Thus colour matches for larger field sizes were not adequately predicted by the 1931 observer. New colour matching experiments were therefore carried out by Stiles and Burch (1959) and by Speranskaya (1959), who obtained colour matching data for a 10° field. To reduce rod intrusion in the colour matching experiments, the luminance of the matching field was kept high. A further correction was made to account for the presence of the Maxwell spot – an area attributed to the macular pigment (see Section 1.6). The macular pigment covers all the retinal receptors in the fovea and beyond and its optical density diminishes sharply with eccentricity. The 10° colour matching functions were derived taking into account the Maxwell spot by masking off a central area. The two sets of experimental results were combined and used to define the 1964 (10°) CIE supplementary standard observer. The supplementary CMFs denoted by $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$, $\bar{z}_{10}(\lambda)$ are shown in superimposed with the 1931 CMFs in Fig. 1.16.

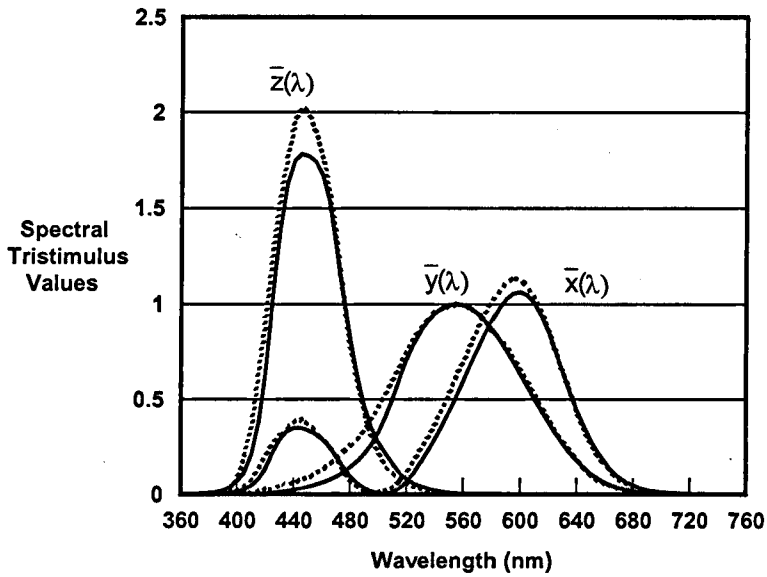


Figure 1.16: The CIE colour matching functions for the 1964 supplementary standard colorimetric observer (dashed lines) superimposed on the 1931 standard colorimetric observer (solid lines). Data obtained from cvrl website (Color vision research laboratories) http://www-cvrl.ucsd.edu/.

The 1931 CIE – (X, Y, Z) system was adopted as the international standard for colour specification and measurement, for use in industry and science. The standard was based on the average colour matching properties of a set of normal trichromats (as there were no reliable estimates of the cone spectral sensitivities at the time). The system meets the requirements of some professionals for precision and objectivity compared to other systems of colour specification such as the Munsell system (in which fading of samples with time causes unreliability). The CIE parameters are based on the spectral power distribution of the light emitted from a coloured object and are factored by sensitivity curves which have been measured for the human eye.

1.5.2 Standard illuminants

Several practical working standards for white lights have been specified by the CIE called standard illuminants. These have been defined based on Planck's radiation formula having a relative spectral power distribution with a certain

colour temperature. The CIE standard illuminants are known as A, B, C and D. Standard illuminant A is representative of the tungsten filament lamp, a common artificial light source with a colour temperature of 2856 K. Standard illuminants representing daylight are of more practical importance than tungsten light however these are more complicated to define. Standard illuminant B (now obsolete) has a correlated colour temperature of 4874 K and was intended to represent sunlight, while illuminant C was intended to represent average daylight (light of the overcast sky for either northern or southern hemisphere) with a colour temperature of 6774 K. More recently, standard illuminant D was defined to represent daylight and is normally written with a subscript where the first two digits indicate the colour temperature. D_{65} is commonly used with a correlated colour temperature of 6504 K representing average daylight throughout the visible spectrum and into the ultra-violet region as far as 300 nm.

1.5.3 CIE chromaticity diagrams

For the CIE 1931 colorimetric system, colour stimuli can be represented on a two-dimensional diagram, with chromaticity coordinates x, y, z given by:

$$x = \frac{X}{(X+Y+Z)}$$
$$y = \frac{Y}{(X+Y+Z)} \qquad (1.2)$$
$$z = \frac{Z}{(X+Y+Z)}$$

From Equation 1.2 it follows that $x+y+z=1$ for all colours. Hence if x and y are known, z can always be calculated as $1-x-y$. It is therefore only necessary to quote two chromaticity coordinates which can be plotted on a Cartesian two-dimensional graph. It can also be shown that X and Z can be easily calculated from x, y and Y; hence this latter set is the acceptable form used for colour specification. By consideration of x, y and Y all colours can be specified. A plot of y against x is called a chromaticity diagram. A similar chromaticity diagram is

derived from the X_{10}, Y_{10}, and Z_{10} tristimulus values for the CIE 1964 standard colorimetric system, by evaluating x_{10}, y_{10}, z_{10} as in Equation 1.2 above.

The chromaticity diagrams associated with the XYZ systems of the CIE 1931 and 1964 standard colorimetric observers are shown in Fig. 1.17.

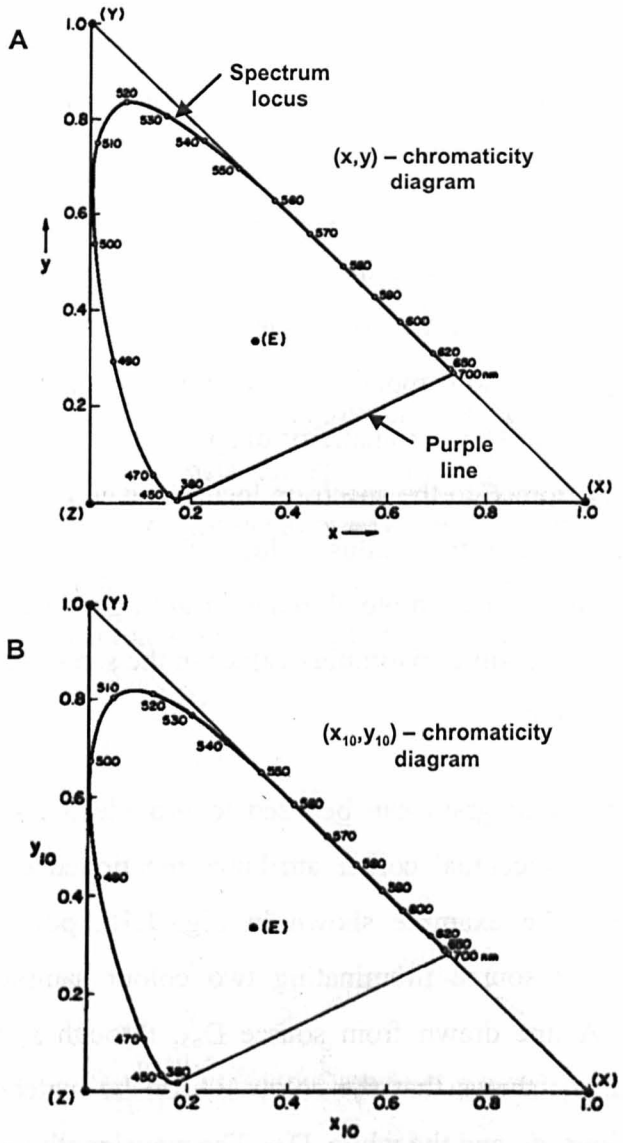


Figure 1.17: The CIE chromaticity diagrams for the XYZ systems of (A) the CIE 1931 and (B) the CIE 1964 standard colorimetric observers. The spectrum locus, purple line and the equal-energy stimulus (E) are shown. Reproduced from Wyszecki and Stiles (1982).

From the definition of the CMFs, the chromaticity point E of the equal-energy stimulus, plots in the centroid of the chromaticity diagram at $x_E = y_E = ⅓$. The chromaticity coordinates of monochromatic stimuli plot as a curve, known as the spectrum locus. The straight line joining the wavelengths at either end of the visible spectrum is known as the purple line, as these colours are mixtures of red and blue. The area enclosed by the spectral locus and the purple line encloses the domain of all colours.

The (x,y)-chromaticity diagram provides a colour map on which the chromaticity of any spectral composition can be plotted. If two stimuli that differ in spectral composition are additively mixed together, then the point representing the mixture is located on the diagram by a point that lies on the line joining the two points representing the two original stimuli. Points that lie close to the spectral locus represent a saturated colour of that particular hue. Contrarily, a point consisting of a mixture mainly of E, corresponding to a point lying near E, will tend to represent a pale colour. The chromaticity diagram thus represents a continuous gradation of stimuli from E to the spectrum locus. The chromaticity diagram only shows proportions of the tristimulus values; thus no information about the luminance is conveyed. For example, dim and bright lights, light and dark surface colours, having corresponding tristimulus values in the same ratios to one another, all plot at the same point.

The CIE chromaticity diagram can be used to provide measures that correlate approximately with perceptual colour attributes mentioned earlier, such as hue and saturation. For the example shown in Fig. 1.18, point D_{65} is a chosen reference white light source illuminating two colour samples represented by points s_1 and s_2. A line drawn from source D_{65}, through s_1 to a point on the spectrum locus at d_1, shows that the colour s_1 can be matched by an additive mixture of the colour, d_1, and the white, D_{65}. The wavelength on the spectral locus corresponding to d_1 is therefore called the 'dominant wavelength' of the colour s_1, relative to the white point, D_{65}. To establish the dominant wavelength of the sample s_2 as the point d_2 is on the purple boundary (has no corresponding wavelength), the line $D_{65}s_2$ is extended in the opposite direction to meet the spectrum locus at point d_{2c}. The corresponding wavelength at this point is called

the 'complementary wavelength' of the colour s_2, relative to the white point, D_{65}. Dominant (and complimentary) wavelength approximately correlates with the hue of a colour.

Saturation of a colour approximately correlates with a term known as 'excitation purity'. For the example shown in Fig. 1.18, excitation purity is specifically a quantity defined by the ratio of the length of the lines joining the specified achromatic stimulus (D_{65}) and the colour stimulus considered (s_1 or s_2), and the specified achromatic stimulus and the monochromatic stimulus on the spectrum locus (through the colour stimulus considered). Therefore excitation purity, p_e, is defined as:

$$p_e = \frac{D_{65}S_1}{D_{65}d_1} \text{ or } \frac{D_{65}S_2}{D_{65}d_2} \qquad (1.3)$$

If the excitation purity has a value of nearly one, the colour is considered highly saturated and will tend to be near the spectrum locus or the purple boundary; conversely if the excitation purity value is near zero, the chromaticity is near the specified achromatic stimulus and the colour will tend to be pale. Excitation purity approximately correlates with the saturation of a colour.

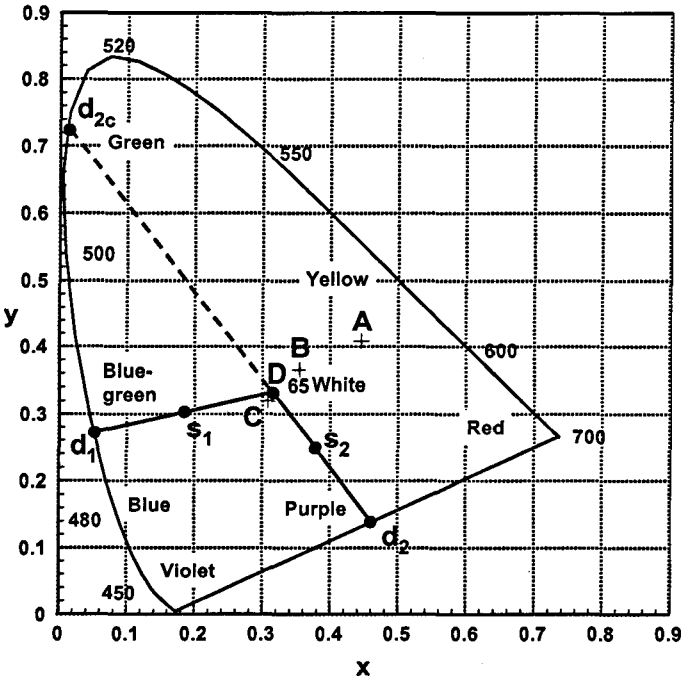


Figure 1.18: The derivation on the 1931 CIE-x,y chromaticity diagram of dominant wavelength, complementary wavelength and excitation purity. s_1 and s_2 are coloured samples and d_1 and d_2 correspond to the dominant wavelengths of these two samples, respectively. The CIE standard sources A, B, C and D_{65} have also been added (see Section 1.5.2). Data obtained from Wyszecki and Stiles (1982).

The chromaticity diagram is useful for specifying colour stimuli and all colours that plot in the same location in this colour space will look exactly the same to a CIE Standard Observer. Two coloured samples with the same (x, y) coordinates or apparent colour have the same tristimulus values no matter what different mixtures of light were used to produce them. The CIE Standard Observer is an average observer based on experiments with small numbers of people (15-20) with normal colour vision. Therefore, the CIE diagram does not define how a colour stimulus will be perceived by a real colour normal observer who is not probably exactly like the CIE standard observer. When two samples match in colour for one observer but fail to match for another observer is known as metamerism. Observer metamerism occurs because of differences in colour vision between observers of biological nature such as individual differences in spectral sensitivity of the photopigments, luminous efficiency function, macular pigment and lens density. Metamerism also occurs between two samples, having different

spectral reflectance factors, matching under one viewing condition but then do not match under a different lighting source.

1.5.4 Chromatic discrimination

It has been estimated that an individual with 'normal' colour vision can in principle discriminate of the order of ten million different colours. This vast number of colours is not solely related to the spectral properties of light, but is a consequence of the very large combinations of just discriminable differences in wavelength, saturation and luminance. The ability of normal trichromats to discriminate between different colour stimuli can be quantified using visual psychophysical techniques. One extensively applied method is the measurement of wavelength discrimination. Two samples of monochromatic light are viewed and compared simultaneously, usually employing a bipartite field, one region of which is filled with light of the fixed test wavelength λ, and the other region is filled with light of variable wavelength. Wavelength discrimination ($\Delta\lambda$) is measured by determining the smallest difference in wavelength that produces a just-noticeable change in appearance between the two halves. The just-noticeable difference between the test and variable fields can be found by direct adjustment whilst maintaining a constant luminosity (according to each particular observer) between the two fields. A wavelength discrimination curve can be obtained by measuring $\Delta\lambda$ for values of λ throughout the visible spectrum. Wavelength discrimination is dependent on field size, eccentricity, presentation duration, luminance level, and stimulus intensity. An example of a wavelength discrimination curve is given in Fig. 1.19.

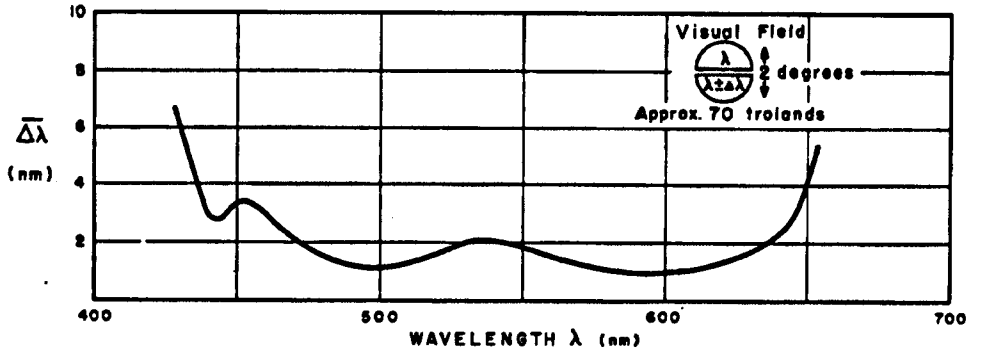


Figure 1.19: Wavelength discrimination curve based on data obtained by Wright and Pitt (1934). The curve is a plot of the average difference $(\Delta\lambda_+ + \Delta\lambda_-)/2$ against λ(nm), where $\Delta\lambda_+$ and $\Delta\lambda_-$ are the just-noticeable differences to longer, and to shorter wavelengths, respectively. Reproduced from Wyszecki and Stiles (1982).

Wavelength discrimination curves characteristically indicate poor discrimination at the ends of the visible spectrum and have two maxima and three minima. The minima represent values at which sensitivity to changes in wavelength is greatest and are in spectral regions where two cone pigments are being stimulated differentially and the dominant sensitivity is changing from one photopigment to another. Their presence is due, in part, to the existence of the three types of retinal cone receptors that have different spectral absorption characteristics.

Chromatic discrimination has also been extended to measurements of just-noticeable differences throughout the chromaticity space. Wright (1941) performed such an investigation using mixtures of pairs of either two monochromatic stimuli, or one monochromatic stimulus and one nonspectral purple. In one half of a 2° bipartite field the mixture of the two lights could be varied, while the other half consisted of a test mixture. Observers were required to adjust the variable field so that it differed from the test field by a constant amount, whilst maintaining a brightness match between the two fields. Wright asked his observers to choose a criterion difference larger than the just-noticeable difference to lessen task difficulty. The results of Wright's experiment are shown in Fig. 1.20. The bars indicate criterion differences in perception, which vary considerably in length in different regions of the 1931 CIE chromaticity diagram.

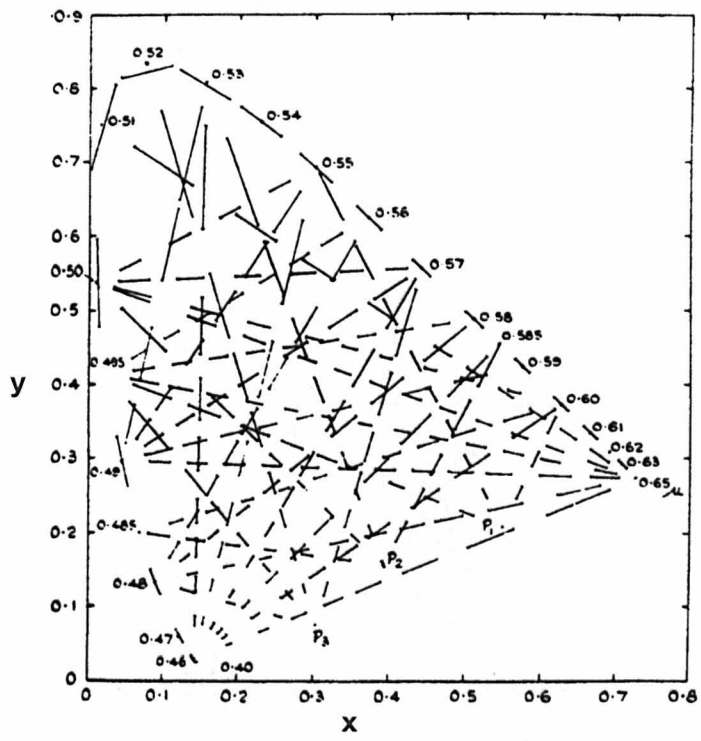


Figure 1.20: The results of Wright's measurements of just-noticeable differences in chromaticity plotted in CIE 1931 chromaticity space. Each line is 3 times the length of a distance representing a difference that is just noticeable in a 2° field. Reproduced from Hunt (1998) after the work of Wright (1941).

MacAdam (1942) represented the chromatic discrimination ability of a normal trichromat by a series of ellipses in the 1931 CIE-x,y chromaticity diagram. The technique of measurement employed by MacAdam involved calculation of the standard deviation of repeated colour matches. In MacAdam's experiment the observer turned a knob that caused the chromaticity in one-half of a bipartite field to vary along a particular direction in CIE 1931 chromaticity space. The fixed stimulus was continuously visible in the other half, which subtended 2° field and was surrounded by a large field of daylight (illuminant C) quality. The luminance of both halves was kept approximately at 50 cd/m^2, and the surrounding field was ~24 cd/m^2. Repeated colour matches were made between the variable field and a fixed stimulus field. The dispersion of a number of such matches was taken as one standard deviation of the distance in the CIE 1931 chromaticity diagram between the variable and the fixed chromaticity point. MacAdam determined in a further experiment that this measure was equivalent to one third of the just-noticeable

difference. For each test chromaticity investigated, MacAdam fitted an ellipse to the discrimination data from different directions of chromaticity space, which were constrained to be symmetrical about the test chromaticity. The results for a fixed chromaticity represented by the point x=0.305, y=0.323 are shown in Fig. 1.21A. An illustration of the parameters used to fit ellipses is shown in Fig. 1.21B.

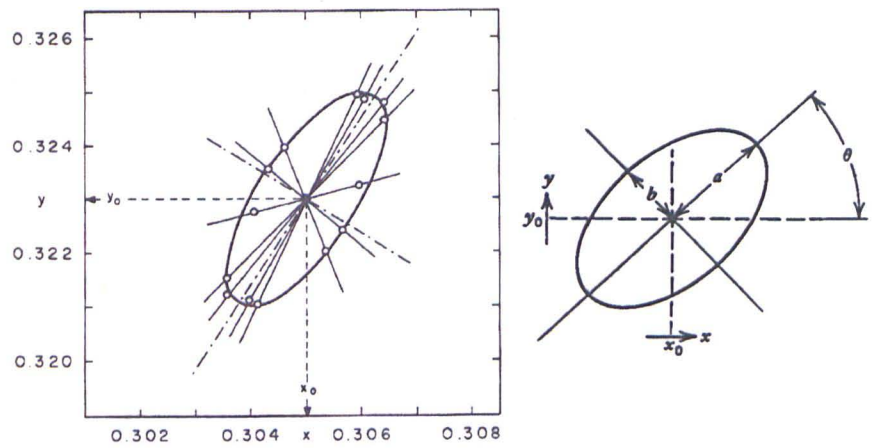


Figure 1.21: (A) An example of the experimental results obtained by MacAdam showing the standard deviations involved in colour matching a fixed chromaticity (x_0, y_0) along different direction in the 1931 CIE–x,y chromaticity space. (B) Illustration of ellipse parameters; (x_0, y_0) = coordinates of centre, a =major semiaxis, b =minor semiaxis, θ =angle in degrees. Reproduced from Wyszecki and Stiles (1982).

A total of 25 locations in chromaticity space were investigated and these ellipses, which have come to be known as MacAdam's chromatic discrimination ellipses, are shown in Fig. 1.22, with axes plotted ten times their actual length. Variations both in the size and orientation of the major axis of the ellipses in the chromaticity diagram demonstrates differences in sensitivity as well as the non-uniformity of the CIE-x,y chromaticity diagram for the prediction of colour discrimination thresholds. The 1931 CIE-x,y diagram therefore does not give a direct way of estimating colour differences. Fig. 1.23 shows that equal distances between chromaticity coordinates of colour stimuli do not correspond to equal visual differences between stimuli.

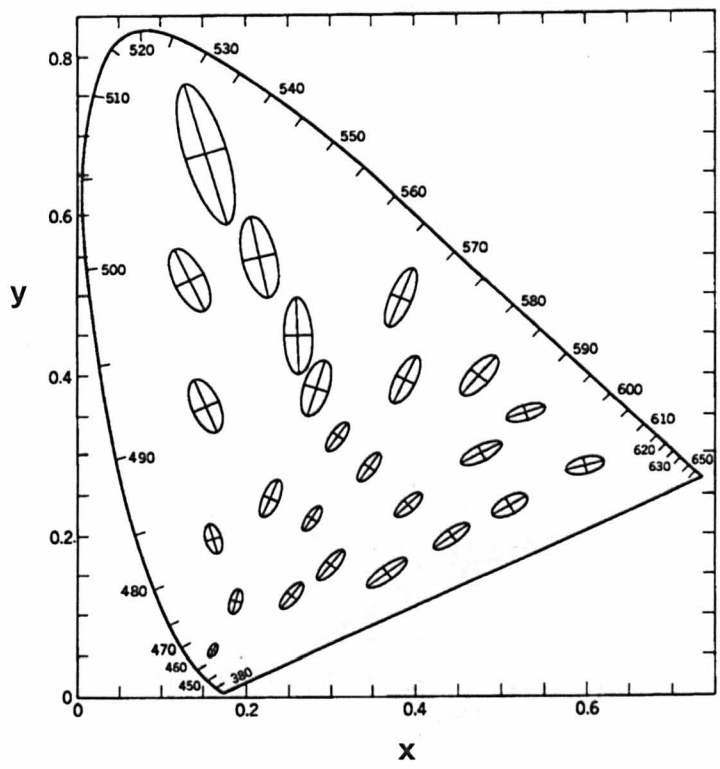


Figure 1.22: The results of MacAdam's measurements of chromatic discrimination based on the dispersion of a series of colour matches, showing MacAdam's ellipses plotted in 1931 CIE-x,y chromaticity space, with ellipse axes 10 times their actual length. Reproduced from Wyszecki and Stiles (1982), after MacAdam (1942).

This lack of perceptual uniformity in the CIE-x,y colour space lead to the attempt to find a mathematical transformation that would produce a perceptually uniform geometry of colour space, i.e., chromatic discrimination ellipses could be represented as circles. In 1976 the CIE recommended the u',v' diagram, but prior to that a similar diagram, the 1960 u,v diagram was developed. The 1976 uniform chromaticity scale diagram or the CIE 1976 UCS diagram is obtained from a conversion of the x,y 1931 CIE coordinates to u',v', by:

$$u' = \frac{4x}{-2x+12y+3}$$
$$v' = \frac{9y}{-2x+12y+3} \qquad \textbf{(1.4)}$$

The advantage of the 1976 diagram is that the distance between points is now approximately proportional to the perceived colour difference (see Fig. 1.23), although not completely, something definitely not true in the 1931 diagram. Despite this obvious advantage, the 1976 diagram is not used as much as the original 1931, which was internationally adopted before any revisions were made.

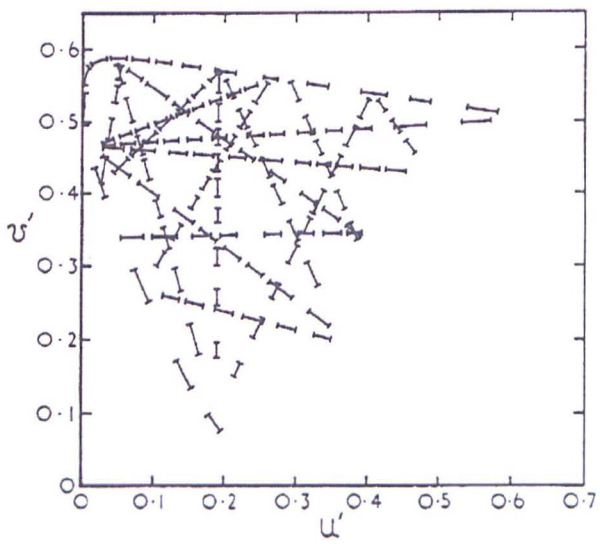


Figure 1.23: The approximately uniform CIE u', v' chromaticity diagram with a selection of the lines (just noticeable differences) of Fig. 1.20. Reproduced from Hunt (1998).

Colour stimuli can also be represented in colour spaces that have physiological significance. Macleod and Boynton (1979) introduced a 2D chromaticity diagram based on isoluminant cone excitations (relative number of quantal absorptions) of the three cone types. Derrington, Krauskopf and Lennie (1984) defined another 3D space where stimuli are modulated along the constant S-cone excitation axis and along a constant L- and M-cone excitation axis with the 3rd dimension representing the achromatic dimension. This is often described as the DKL.

1.6 Individual differences in colour vision

More than 95% of people exhibit nearly the same responses when matching and identifying colours in objective tests, so there seems to be a form of normal trichromatic vision that most of us share. Trichromacy is based on the existence of three cone types sensitive to different parts of the visible spectrum. However not everyone has trichromatic vision and many people are partially colour deficient, confusing colours that full trichromats distinguish. In the general population, colour vision deficiencies can be acquired or of congenital origin, the latter being much more common.

1.6.1 Variation in 'normal' colour vision

Individual variations in colour vision occur as a result of individual differences in pre-receptoral and receptoral absorption properties. Physiological differences that may cause variation in the spectral absorption properties of the retinal photoreceptors are the optical density and peak sensitivity of the cone photopigments. The optical density of the photopigments cause a broadening of the spectral absorption curves for higher densities and the variation in the peak absorbance of cones cause differences in the relative sensitivity at different wavelengths. Further, the relative numbers of L- and M-cone photoreceptors are likely to affect the individual spectral luminous efficiency functions of different eyes, but may not affect significantly colour vision (Kremers et al. 2000). Inter-observer variations therefore occur due to differences in photoreceptor pigments and thus in the photopic luminosity curves, since the later is dependent on signals from the cone photoreceptors.

In front of the retina there are a number of spectrally selective filters affecting the light reaching the visual photopigments. The absorption of light above 500 nm is likely to be affected by the water content of the ocular media (humours). For wavelengths shorter than 500 nm, water is transparent to electromagnetic energy but absorption does occur due to the other components of the media, such as the cornea and the lens. The lens is the principal medium which appreciably affects

the spectral distribution of incident visible light (380-780 nm).The lens prevents any wavelengths shorter than 300 nm from reaching the retina, having a high optical density at 400 nm. The transmission properties of the lens in the eye are not uniform across all wavelengths as shown in Fig. 1.24. The optical density function decreases rapidly above 450 nm, and the lens transmits over 90% of incident light for wavelengths longer than 580 nm (Van Norren and Vos 1974). During life, changes in the lens are accompanied by increases in absorbance of the lens (Pokorny et al. 1987; Weale 1988). The light losses of the eye lens are markedly dependent on the age of the eye (see Fig. 1.24). The reduction in transmission is due not only to an increase of the absorbance but also to an increase in the scattering of light in the lens which also has a greater effect on the shorter wavelengths.

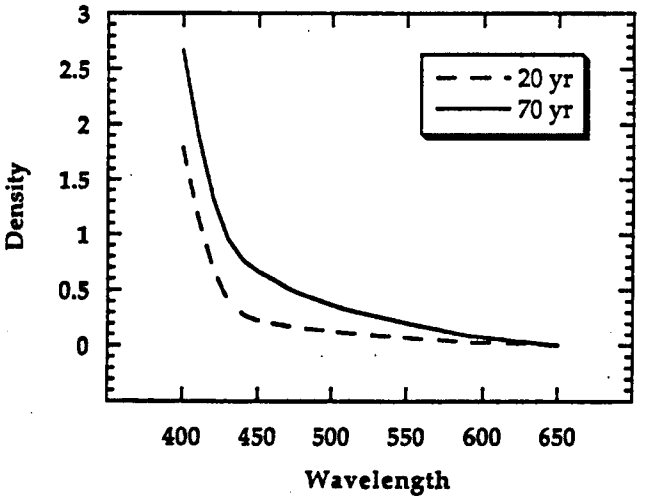


Figure 1.24: The average optical density spectrum of the lens absorption of a 20 year old (dashed line) (Van Norren and Vos 1974), and that of a 70 year old (solid line) (Pokorny et al. 1987). Reproduced from Pokorny and Smith (1997)

The macular pigment also acts as a selective filter modifying the spectral distribution of incident light. It is found in the macular area of the retina just in front of the photoreceptors. The absorption characteristics of this pigment are very similar to that of carotenoid pigments. It is yellow in appearance and absorbs light preferentially in the blue region of the visible spectrum with peak absorption at

approximately 460 nm (Bone et al. 1992) as shown in Fig. 1.25. Macular pigment has been estimated from the differences between measurements of absolute threshold obtained at the fovea and in the periphery where there is little or no density of pigment, or from the differences between foveal and peripheral colour matches (Ruddock 1963). Individual variations as high as 1 log unit optical density have been measured (Ruddock 1963).

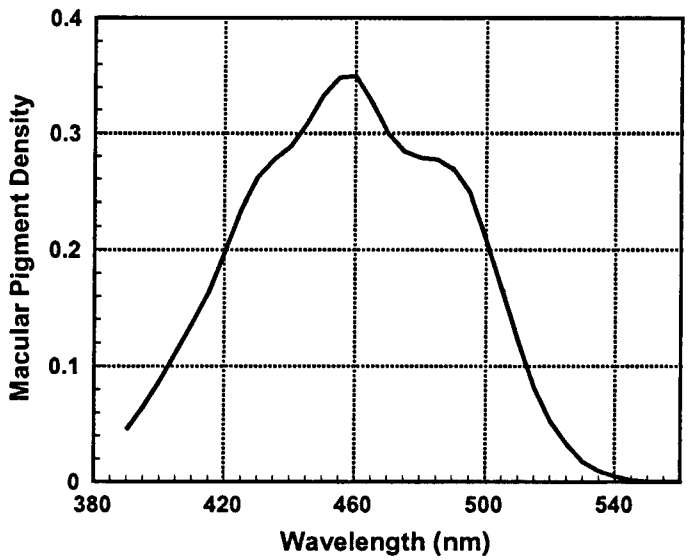


Figure 1.25: Absorption spectrum of the macular pigment (Bone et. al. 1992). Data obtained from cvrl website (Color vision research laboratories) http://www-cvrl.ucsd.edu/.

Among colour normals, there are also several conditions that cause trichromacy to collapse. As night falls, colours fade and become paler merging into shades of grey; this is the transition from cone- to rod-mediated night vision. Further, a reduction in yellow-blue colour vision also occurs: when viewing small targets centrally (foveal tritanopia) where S-cones are missing (Curcio et al. 1991); and when viewing small targets in the far periphery (small field tritanopia), where S-cones are sparse; and when viewing immediately after the extinction of strong, yellow adapting fields (transient tritanopia). Under extreme conditions of chromatic adaptation and photopigment bleaching, trichromacy may also fail to a certain extent.

In some circumstances the absorption of light in the ocular media can have an important effect on the resultant colour perception. Both the variation of the lens absorbance (with age) and the absorption of the macular pigment (between individuals) must be taken into account when comparing the colour vision of different individuals.

1.6.2 Congenital colour vision defects

The majority of the population exhibit normal trichromatic vision, yet there is a distinct group of the population with deviant colour vision that is clearly not a result of physiological or perceptual variations in a single population group. Statistically they comprise a small group of individuals who are defined as having defective colour vision. There are two causes of defective colour vision; either inherited or acquired as a consequence of pathology. Congenital colour vision deficiencies are more common amongst men, ~8%, with an incidence of only ~0.4% in women (see Table 1.1).

Type of colour deficiency	Male (%)	Female (%)
Protanopia	1.01	0.02
Protanomalous trichromatism	1.08	0.03
Deuteranopia	1.27	0.01
Deuteranomalous trichromatism	4.63	0.36
Tritan defects	rare	rare
Number	45.989	30.711
Total prevalence (%)	7.99	0.42

Table 1.1: Incidences of colour vision deficiencies in men and women. The raw numbers were taken from several different European populations. The table shows that protanopia and deuteranopia are equally frequent in the European male population, whilst deuteranomaly is four times more frequent than protanomaly. The incidence of tritan defects is very rare. Data obtained from Sharpe et al. (1999).

Congenital colour vision deficiencies arise due to alterations in the genes encoding the molecules defining the photopigment. The spectral sensitivity of one of the three photopigments groups can differ significantly from that of the normal, or one/two of the three photopigment groups are absent or rendered non-functional. These different types of colour deficiency are classified according to the number of photopigments present and hence the number of colour matching variables required to match all the spectral hues. Visual acuity and all other visual functions are not affected by congenital colour deficiency. The deficiency is binocular, symmetrical and stable over time.

Monochromacy is a form of total colour blindness. Monochromats often claim to perceive or associate colours with objects by discerning differences in their brightness. It occurs when two or all three of the cone pigments are missing and, colour and luminance vision is reduced to one dimension (complete achromatopsia).

Dichromacy occurs when one class of cone photopigment is missing and colour vision is reduced to two dimensions (as opposed to three required by normal trichromats). There are three classes of dichromatism depending on which of the three normal pigment types is functionally absent. Protanopia and deuteranopia refer to the condition of loss or absence of L- and M-cone pigments, respectively, whereas tritanopia refers to the lack of S-cone pigments. The most common forms of dichromatism are deuteranopia and protanopia. Very few females have this deficiency.

Anomalous trichromacy is the condition when all three types of cone photopigments are present, but one of the three cone pigments is altered in its spectral sensitivity. Protanomaly, deuteranomaly and tritanomaly refers to the presence of abnormal L-, M- and S-cones, respectively. Deuteranomaly is the most common form of anomalous trichromatism while tritanomaly is very rare.

1.6.2.1 Incidence of protan and deutan defects

The incidences in males are much higher than in females because the defects are inherited as recessive traits found on the X-chromosome, which is also the chromosome linked to sex determination. In humans, the DNA sequence is made up of 23 pairs of chromosomes, one pair concerned with sex determination. In males, sex chromosomes are dissimilar one labelled X and one labelled Y; in females the two sex chromosomes are similar and labelled XX. An X-linked recessive gene carried on the X-chromosome is not manifest in the female in the heterozygous state. Heterozygous females are described as carriers of X-linked traits. Males, who have only one X-chromosome, are hemizygous and will always manifest the colour defect if they inherit an aberrant gene on the single X-chromosome and there is no matching gene on the Y chromosome. Females, on the other hand, will only manifest the colour vision defect if they are homozygous (identical genes). Fig. 1.26 shows a schematic of how typically X-linked chromosome defects are transmitted.

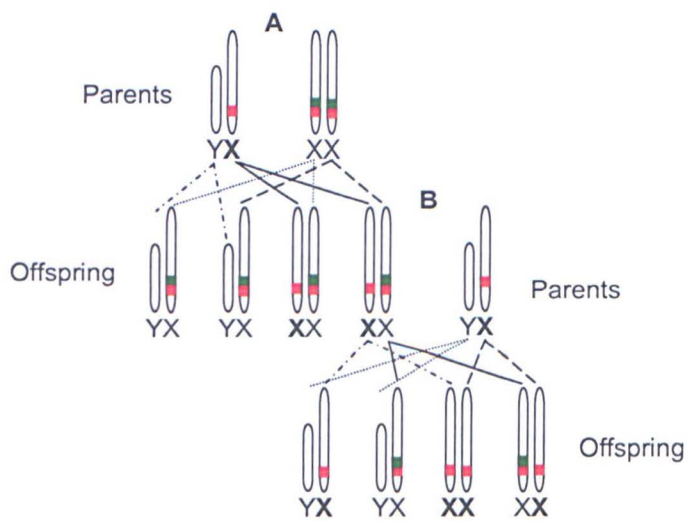


Figure 1.26: Schematic of the X-linked recessive inheritance for deutan (or protan) defects. Red and green squares indicate L- and M-cone, respectively. (A) In the example shown the father (affected) lacks the M-cone opsin gene on his single X-chromosome and the mother has normal colour vision. The descendent lines show that the boys will have normal colour vision, while the girls will be carriers. (B) If now the mother lacks an M-cone (carrier) on one of the X-chromosomes and the father also lacks the M-cone opsin gene on his single X-chromosome. The descendant lines show that on average, half of the boys will be affected hemizygous and the other half normal; half of the girls will be carriers (heterozygous) and the other half affected (homozygous).

The prevalence of X-chromosome-linked or red/green defects, as they are commonly known in the human population, is shown in Table 1.1. Large surveys show that protanopia, protanomalous trichromatism and deuteranopia each occur in approximately 1% of men and that deuteranomalous trichromatism occurs in about 5% of men. Some studies have also shown variability between human populations of different racial origin. The highest rates are found in Europeans and Brahmins of India; the lowest in the aborigines in Australia, Brazil, the South Pacific Islands, and North America (Sharpe et al. 1999).

A characteristic of X-linked inheritance is that the affected family members have the same type and severity of colour deficiency. From the typical prevalence of colour deficiency in men which is of the order of 8%, the prevalence of colour deficiency in women can be calculated and should be (0.08^2) 0.64%. Table 1.1 shows that in practice there is a much lower prevalence of around 0.4%. About 15% of women are carriers of an abnormal opsin gene of red/green colour vision defects. In heterozygous women, X-chromosome inactivation means that one of the two chromosomes is transcriptionally silenced. This process is normally random, with an equal probability that either the maternal or paternal inherited X-chromosome is inactivated. Women with two abnormal genes coding for the colour deficiency, in the heterozygous state one gene is dominant and the other recessive; the person manifests the characteristics of the dominant gene. However women with mixed abnormal genes, one abnormal gene coding for protan defect on one X-chromosome and an abnormal gene coding for deutan defect on the other, can manifest normal colour vision (Jordan and Mollon 1993a) and this explains the lower prevalence observed.

1.6.2.2 Incidence of tritan defects

As for protan and deutan defects, congenital tritanopia arises from alterations in the gene encoding the opsin, but, unlike the protan and deutan defects, it is autosomal in nature. In autosomal dominant inheritance, the individual inheriting a defective gene will show the defect, and the offspring will have affected

members in each generation (see Fig. 1.27). Autosomal dominant defects are characterised by variable expression and severity of the abnormality from one family member to another (Cole et al. 1966). The frequency of tritan defects should be equivalent in males and females, but the actual incidences have never been accurately established. Due to the relatively low number of cases, it is common to use the all-inclusive term tritan defects to include tritanopia and tritanomaly. Identification of slight tritan defects may be within the normal variations in macular pigment density and in age-related lens density. Estimated rates of incidence in the UK are 1 in 13,000 to 1 in 65,000 (Wright 1952). A study conducted in the Netherlands suggests that the prevalence of tritan defects to be as high as 1:500 (van Heel et al. 1980; van and Went 1981).

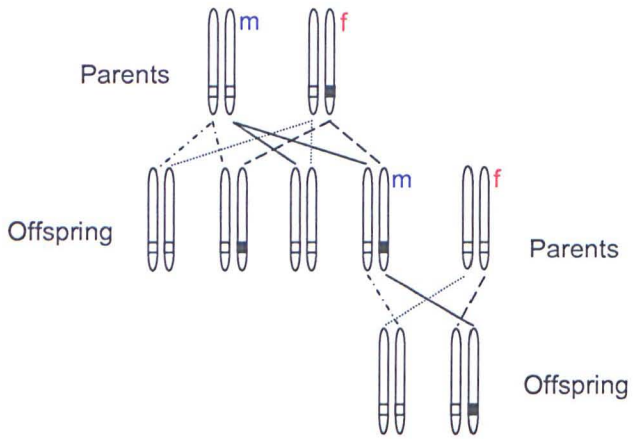


Figure 1.27: Autosomal dominant inheritance. On average, one-half of the offspring of a heterozygotic parent will inherit the disorder. An open square indicates a normal functioning S-cone pigment gene; a filled square, one carrying a mutation perturbing the structure and stability of the S-cone pigment.

1.6.2.3 Phenotypes of congenital colour vision deficiencies

The special terms used for defective colour vision (protan, deutan and tritan) are derived from the Greek words protos, deuteros and tritos meaning first, second and third, respectively. This is a purely nominal classification based upon the

order in which the defects were originally discovered and gives no indication about severity or prevalence of the different types of defective colour vision.

A number of non-invasive colour vision tests have been designed which exploit the colour confusions of colour deficient observers. These include colour confusion charts, e.g., the Ishihara pseudoisochromatic plates, American Optical Corporation Hardy, Hand and Ritler plates, hue discrimination or arrangement tasks: Farnsworth Panel D-15, and lantern detection tests e.g., the Holmes-Wright (Rayner, England), Beyne (Luneau, France). Traditionally, observers with protan and deutan defects are characterised by special colour matches employing only two primaries, performed on visual colorimeters known as anomaloscopes. Several of the above tests will be described in Chapter 2.

There are many ways in which the perceptual characteristics of colour vision show systematic differences amongst the several types of colour vision deficiency. These include spectral sensitivity, wavelength discrimination, chromatic discrimination, and colour matching ability.

Any change in the absorbance spectra of the cone photopigments (Alpern 1979; Pokorny et al. 1973; Pokorny and Smith 1977; Shevell and He 1997a) results in a change of the photopic spectral sensitivity of the visual system. The relative luminous efficiency shows characteristic properties for colour deficient observers which enables to identify the affected or absent photopigment. The three types of dichromats also differ in which part of the equal-energy spectrum appears brightest (Wyszecki & Stiles 1982). For a normal, the maximum of the luminosity function is on average near 555 nm (see V(λ) curve in Fig. 1.15), for protanopes it is around 540 nm, and for deuteranopes closer to 560 nm (Pitt 1935). Tritanopes have the same peak sensitivity as normals, but their relative luminous efficiency functions display reduced sensitivity at short wavelengths (Wright 1952). Examples of luminous efficiency curves for dichromatic observers are illustrated in Fig. 1.28.

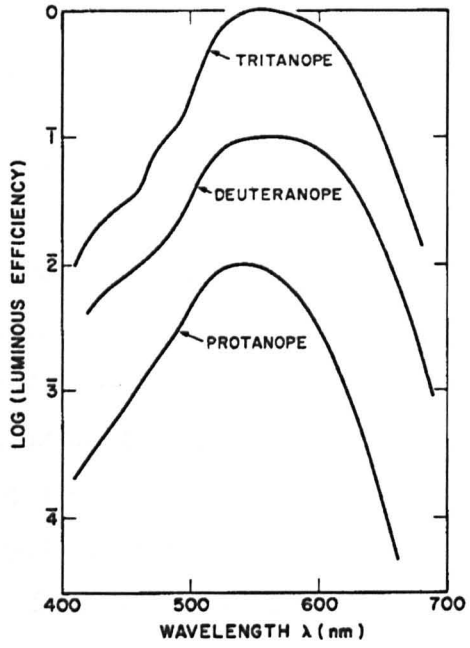


Figure 1.28: Average luminous efficiency functions for protanopes (Pitt 1935), deuteranopes (Pitt 1935) and tritanopes (Wright 1952). The curves are displaced vertically for clarity. Reproduced from Wyszecki and Stiles (1982).

In protanopia, the absence of the L-wavelength sensitive cones means that there is a considerable reduction in sensitivity at the red end of the spectrum. For example, stimuli typically perceived as bright red to the normal trichromat will appear very dull and dark to the protanope. This provides an additional valuable distinction between protanopia and deuteranopia. However, for deuteranopes, the luminous efficiency function is derived from the absorbance spectra of the S- and L-cones so, due to the similarity of the shapes of the spectra for the L- and M-cones on the short wavelength side of their peaks, its peak sensitivity and overall shape is not that different from that of the normal trichromat. The relative luminous efficiency functions of deuteranomalous and tritanomalous subjects obtained using flicker photometry, are similar to those of normal trichromats, whereas protanomalous, like protanopes, have insensitivity to the far end of the visible spectrum, with a blue-shifted luminosity function (Wright 1946).

The wavelength discrimination function (see Fig. 1.19 for normals) for protanopes and deuteranopes is U-shaped with a single minimum near 500 nm. Protanopes and deuteranopes can discriminate from 440-520 nm and 430-530 nm,

respectively. Wavelength discrimination for tritanopes is absent from about 445 to 480 nm. For anomalous trichromats, wavelength discrimination covers a fairly continuous range from nearly normal to almost dichromatic (Wright 1946).

Colour discrimination by dichromats is limited due to one photopigment being absent. Therefore, when it comes to colour matching, certain colours that are easily distinguished by normal trichromats will appear identical to dichromats (if there is no luminance contrast), and are therefore confused. These confusions can be represented in the 1931 CIE-x,y chromaticity diagram (see Section 1.5.3), as the MacAdam ellipses for normal trichromats are (see Fig. 1.22), but the major axis of each discrimination ellipse extends the full length or width of the chromaticity diagram and is typically depicted as a line (Maxwell 1855). The lines shown on Fig. 1.29 are known as isochromatic lines as they describe the locus of chromaticities which all appear as the same colour, or simply as confusion lines as they show the colours that are confused by dichromats. In anomalous trichromacy the chromaticity discrimination thresholds are represented by elongated MacAdam ellipses with major axes (that do not extend the full length) corresponding to the direction of relevant isoconfusion lines.

All confusion lines converge to a point called the copunctal point, which have been measured experimentally and are located outside the spectrum locus (Smith & Pokorny 1975). Protanopes, deuteranopes and tritanopes are distinguished because the isochromatic lines lead to different copunctal points, corresponding to the chromaticity of the L-, M- and S-cones, respectively (see Fig. 1.29). A confusion line often used in clinical colour vision tests is one that goes through the neutral point or achromatic stimulus (equal-energy). The neutral points for dichromats lie at the intersection with the spectrum locus of a straight line through the copunctal point and the particular achromatic stimulus. For protanopes and deuteranopes there is only one neutral point, but for tritanopes there may be two intercepts of the specified straight line with the spectrum locus (Wyszecki & Stiles 1982).

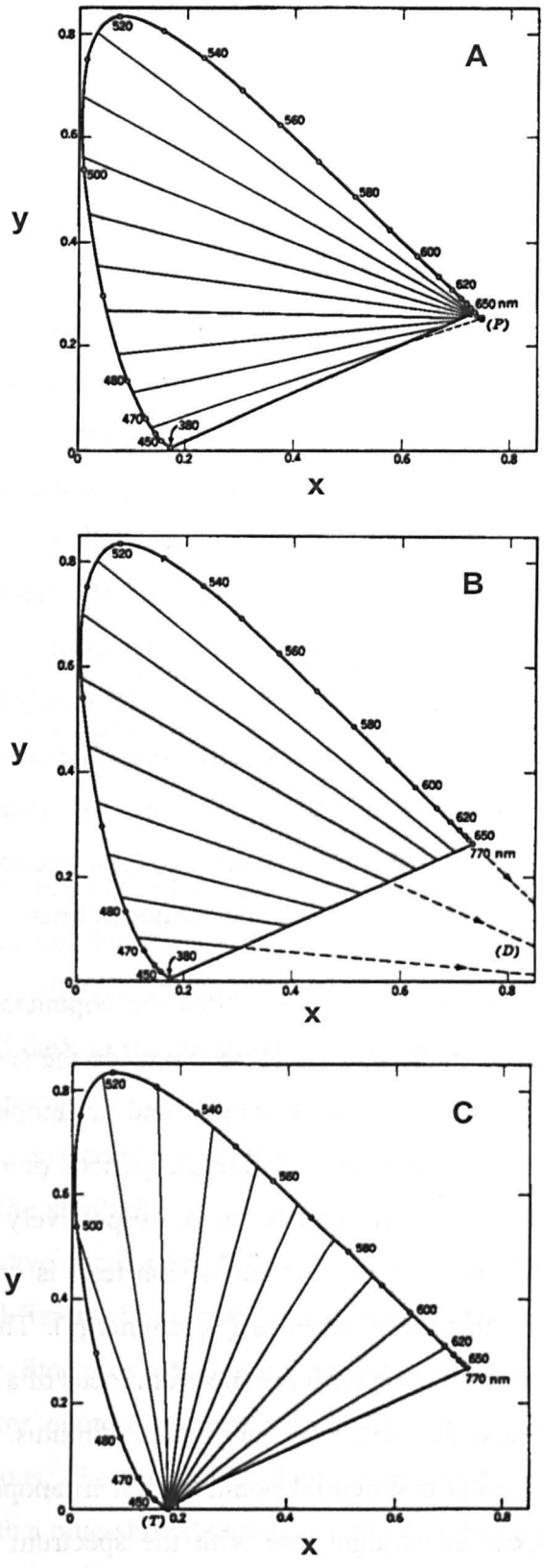


Figure 1.29: Dichromatic lines of constant chromaticity or hue for (A) protanopes, (B) deuteranopes, and (C) tritanopes, plotted on the normal trichromat's 1931 CIE-x,y chromaticity diagram. The lines have been extrapolated to converge at the confusion or copunctual point for (P) protanopes, (D) deuteranopes (falls outside the graph), and (T) tritanopes. Reproduced from Wyszecki and Stiles (1982).

The colour vision defects of anomalous trichromats are usually less severe than those of dichromats. Three colour matching variables are required to match all spectral hues (as for normal trichromats). A feature of the matching behaviour of protanomalous and deuteranomalous trichromats is considerable variability among observers. The variability is so great that the data for individual observers have not been averaged to give mean sets of colour matching functions analogous to that available for normal observers. Many simple anomalous trichromats may be unaware of their colour vision deficiency, whereas many extreme anomalous trichromats exhibit wide matching ranges with a loss of chromatic discrimination similar to that of dichromats.

These findings are consistent with the large individual variation found in the genes encoding the anomalous photopigments (Asenjo et al. 1994; Merbs and Nathans 1992b; Nathans et al. 1986a) and in the wavelength of peak sensitivity of individuals' cone photopigments (Alpern 1979; He and Shevell 1995; Pokorny & Smith 1977; Shevell & He 1997a).

1.6.2.4 Monochromacy

Monochromats can match all spectral wavelengths using only one colour-matching variable. There are two principle types of monochromacy: when receptors are either absent or non-functional, the condition is described as rod monochromacy (typical) and when there is evidence of some cone function and people have normal visual acuity, it is known as cone monochromacy (atypical).

Rod monochromacy is a rare congenital disorder, which is inherited as an autosomal recessive trait, also referred to as typical, complete achromatopsia, sometimes accompanied by reduced visual acuity and photophobia (aversion to bright light), hyperopia (farsightedness), nystagmus (unsteady fixation), and the complete inability to discriminate between colours. Manifestation is functionally equivalent to the total loss of all three cone opsin genes. Although the opsin genes may be expressed, they are not involved in vision. The prevalence of typical rod

monochromacy found in a Northern European survey in 1990 produced the figure of 1 in 50,000 for both sexes (Sharpe and Nordby 1990).

There are two main classes of cone monochromacy that may occur: S-cone monochromacy and L-/M-cone monochromacy. In S-cone monochromacy, only the S-cones are believed to be functioning, therefore the perception of colours is non-existent. However, there are a number of reports of S-cone monochromats showing some residual colour vision. This is thought to be possible due to interactions between the S-cones and rods at mesopic levels where both rods and cones are active (Alpern et al. 1971; Pokorny et al. 1970). This disorder is characterised by severe reduced visual acuity often accompanied by nystagmus and photophobia. The photopic luminous efficiency function peaks near the maximum wavelength (λ_{max}) of the S-cone spectral sensitivity function (see Fig. 1.28). The estimated prevalence is about 1 in 100,000 and all the cases reported so far have been male (Sharpe et al. 1999).

The other forms of cone-monochromacy are also known as complete achromatopsia with normal visual acuity. Again only a few cases have ever been described which vary from 1 in 1,000,000 to 100,000,000 (Pitt 1944; Weale 1953). These are conventionally known as L- or M-cone monochromacy if they have either L- or M-cones, respectively, but not both. Unlike S-cone or rod monochromacy, there is no reduced visual acuity, nystagmus or light aversion. There is some evidence of remnant cone function which has led to the speculation that the defect may be somehow postreceptoral in origin (Weale 1953).

1.6.3 Acquired colour vision defects

Several aspects of human vision such as visual acuity, the extent of the visual field, and colour discrimination improve during childhood and adolescence, achieve their maximum values in the third decade, and thereafter deteriorate progressively with chronological age. Gradual deterioration, in particular of the yellow-blue channel, has been demonstrated by subjects' performance with several colour vision tests (Knoblauch 1987; Lakowski 1958). However, a more

recent study has shown similar effects on chromatic discrimination due to advancing age for both the yellow-blue and red-green channels (Knoblauch et al. 2001). These senile changes in colour vision are mainly due to the progressive yellowing of the crystalline lens (see Fig. 1.24).

There are less common forms of colour vision deficiency that arise from factors other than inherited alterations in the opsin genes. These are usually progressive and tend to be associated with other related signs, such as: disorders of the pre-receptoral ocular media, retinal detachment, progressive cone dystrophies or degenerations affecting all cone classes, macular dystrophies and degenerations, vascular and hematologic diseases, glaucoma (Pacheco-Cutillas et al. 1999), hereditary optic atrophy and other optic nerve diseases, diseases of the central nervous system (e.g. multiple sclerosis), or other organs (e.g. diabetes), toxic agents (e.g. tobacco, lead, alcohol) that affect the retina or optic tracts. A full review on acquired colour vision defects was produced by Birch (1979).

2 Methods of Assessing Colour Vision

The increased interest in different methods for evaluating human colour vision during the last century is usually attributed to the development of mass transport with the accompanying mandate to screen colour defective observers from sensitive occupations in industry. Clinical testing of congenital colour deficiencies is designed for easy and rapid assessment of colour vision. Laboratory measurements of photopigment absorption spectra allow for advanced analysis of colour vision characteristics. However, such measurements require specialised equipment, take a long time to collect and need a trained observer.

In this chapter, we start by describing the most important colour vision tests. Emphasis is placed on the 'battery' of colour vision tests employed in this work and the key parameters that are used to describe and evaluate the outcome of such tests. We then examine the criteria defining the normal observer on the gold standard reference test using subject data measured as part of this work. Finally, we present a comparison of the different tests versus the standard reference test.

2.1 Introduction

Clinically, the assessment of colour vision is usually performed to screen from the normal population those individuals who have a congenital red/green or yellow/blue colour vision defect. One often requires to identify and differentiate congenital and acquired colour vision deficiencies. In some working environments one needs to select personnel with normal colour vision. Some of these tests replicate, as far as possible, the conditions of observations in particular jobs, e.g., the lantern tests used for admission to the Royal Navy, railway, etc., in which the candidate is asked to name colours of 'point' light sources that are presented. Most clinical colour vision tests, however, do not replicate specific occupational conditions but are designed on more general lines. Nearly all rely on the poor discrimination of the colour-deficient observer, which leads to mistakes in colour naming, colour matching, ordering of colours and the recognition of coloured patterns.

Various techniques have been exploited to assess colour vision or the loss of chromatic discrimination sensitivity. Colour deficient people, for example, have difficulty finding an object in everyday life when that object is defined by colour but lies in a variegated background where lightness is differing randomly, such as fruits or berries amongst foliage (Steward and Cole 1989). The first colour vision tests, developed in the 1870s, attempted to reproduce this situation by designing a solid figure of one chromaticity on a uniform background of another chromaticity. However, it was immediately realised that it was impossible to create a figure and a background in such a manner as to eliminate all edge artefacts. Further, the equally light figure and background for one observer would not be necessarily the same lightness for another colour deficient observer. In 1877, it was Stilling (see Regan et al. 1994) who redressed these two problems by breaking up both target and background stimuli into a number of discrete patches, each with its own limiting contour and of varying lightness. Only the patches that make up the figure will have the same chromaticity, none of the background patches will share chromaticities belonging to the figure. Thus, neither edge artefacts, nor luminance differences could be used as a cue for discrimination of the target against the background field.

Further approaches towards the assessment of colour vision involve variations in colour matching ability and measuring chromatic discrimination thresholds. These techniques are based on the characteristic colour confusions found in protan, deutan and tritan defects (see Section 1.6.2) (Pitt 1935; Wright 1952). With reference to the characteristic isoconfusion lines of colour vision deficiencies, as represented in the 1931 CIE-x,y chromaticity diagram (see Fig. 1.30), it is possible to select pairs of stimuli whose chromaticities will be discriminated by individuals with normal colour vision yet appear indistinguishable selectively to protan, deutan or tritan colour deficient observers.

Colour vision tests are designed to perform different functions (Birch 2001): some tests are used for screening or identifying people with anomalous colour vision; some tests classify the type of colour deficiency (i.e., protan, deutan and tritan); grading tests estimate the severity of colour deficiency; and, practical colour matching ability, hue discrimination and colour recognition are examined for occupational suitability.

A wide variety of practical tests is available that differ both in design principles and method of administration. One group of colour vision tests is based largely on the principle of colour confusion and tends to be the simplest to use. The majority of such tests are known as pseudo-isochromatic tests, as they employ stimuli that may be described as falsely appearing of the same colour (e.g. Ishihara plates). These are usually presented in book form and the verbal identification of a coloured figure (of specified chromaticity viewed against a background of different, but carefully selected, chromaticity) is required. More sophisticated judgement is needed for both hue discrimination tests, designed principally to assess colour discrimination, and red/green colour matching tests, based on the principles of matching a reference stimulus by a mixture of light intensities from two spectral primaries. Other tests are more specialised, designed to measure specific aspects of colour vision and overall performance, such as judgements of colour difference and colour memory.

There is no international consensus on a standard procedure for examination of colour vision and clinical assessment often relies on the use of several different

tests that often produce inconsistent results (Squire et al. 2005) which have to be combined for reliable clinical diagnosis.

2.2 Clinical and occupational colour vision tests

2.2.1 Test Battery

To provide a detailed description of a person's colour vision, the use of several different tests is often required. This is known as a test battery. The tests included in a battery may be limited by the availability, location, cost, examiner's experience and the time available. A number of batteries have been proposed for the evaluation of congenital colour vision defects (Verriest 1968) and acquired colour vision defects (Pinckers and Baron 1978).

The test battery used in this study is outlined in Table 2.1.

Colour Vision test			
Test	**Type of test**	**Version**	**Use**
The Ishihara plates	Pseudoisochromatic plates	38-plate test (first 25 used)	Screening
The AO-HRR plates	Pseudoisochromatic plates	2nd ed.	Grading & classifying
The Farnsworth-D15	Hue discrimination		Grading & classifying
The CU test	Hue discrimination	2nd ed.	Grading & classifying
Nagel Anomaloscope	Spectral anomaloscope	Type I	Grading, classifying & diagnosis
CAD test	CRT monitor, chromatic discrimination		Grading, classifying & diagnosis

Table 2.1: Test battery carried out at City University.

The first five colour vision tests will be described in the following sections and the new CAD test will be presented in Chapter 3.

2.2.2 Plate tests

There are many different plate tests available, but all tend to reproduce similar aims and formats. The main purpose of plate tests is to screen congenital protan and deutan colour defective observers from normals, some attempt to provide some diagnosis of both type and severity of defect, and some contain plates designed to identify tritan defects.

Most of the plate tests are based on the pseudoisochromatic principle; the use of colour camouflage by carefully chosen colours based on the isochromatic confusion lines (Pokorny et al. 1979). Spots or patches of colour that vary in size and lightness make up the individual elements of the designs in such a way that a different figure emerges from the background for people with normal colour vision to that seen by colour deficient people. The colours contained in the figure and background are within isochromatic zones, so that all the spots appear 'falsely of the same colour' to colour deficient observers. The outline of the figure is broken up and its shape concealed by the dot matrix.

The efficiency of each plate depends on the accurate choice of colour, and positioning of the dots, the dimensions of the figure, and the luminance contrast between the elements contained in the figure and those in the background (Birch 1975). The colour differences should be carefully selected, along isochromatic lines, small enough so that slight defects can be detected and normal subjects can see the figure. It is also crucial to control the luminance contrast between the different elements of the figure and background. If residual luminance contrast signals remain unmasked, the figure may be distinguished using achromatic contrast alone. A variety of colour tones are used specifically to reduce the luminance contrast and to allow for inter-observer differences in relative luminous efficiency (see Section 1.6.2). Isochromatic zone colours for protans and deutans are used for the spots on screening plates for identifying red/green colour deficiency. The design of pseudoisochromatic plates for yellow/blue defects is particularly difficult due to large variations among normal observers in luminous efficiency for the short-wavelengths of the visible spectrum and differences in pre-receptor filtering (see Section 1.6.1). Further, due to the rareness of

yellow/blue colour defects, no adequate detailed information concerning the typical colour confusions of yellow/blue colour defectives and differences between congenital and acquired yellow/blue defects is available.

2.2.2.1 Ishihara pseudoisochromatic plates

The Ishihara test is the most widely accepted screening test for congenital protan and deutan defects. It was first published in 1917 and since then has been reprinted in many different editions, including abridged versions and is used worldwide. Fig. 2.1 shows three of the Ishihara plates. Several clinical trials (Belcher et al. 1958; Birch 1997c; Frey 1958; Sloan and Habel 1956) have shown that the Ishihara pseudoisochromatic test is a highly effective screening test. The accuracy in the test design in terms of compliance with isochromatic data for red/green colour deficiency was established by Lakowski (1965a).

The 38-plate edition has been recommended for clinical use (Birch and McKeever 1993). In this edition the first 25 plates contain numerals (single or double-digit numbers) which have to be identified verbally. The remaining plates have pathway designs intended for the examination of non-verbal subjects and are rarely used. The Ishihara test employs a range of designs that can be divided into five categories (Pokorny et al. 1979):

1. Introductory (or Demonstration) - plate no.: 1
2. Transformation (or Confusion) - plate nos.: 2-9
3. Vanishing - plate nos.: 10-17
4. Hidden Digit - plate nos.: 18-21
5. Classification - plate nos.: 22-25

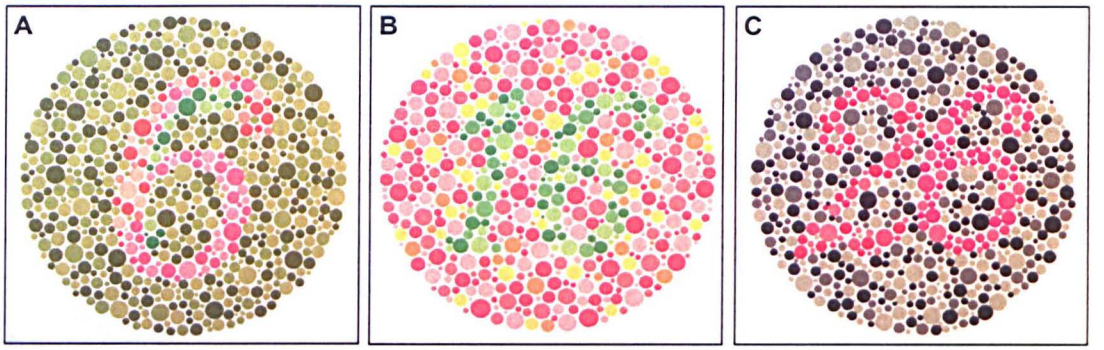


Figure 2.1: Ishihara pseudochromatic plates showing: (A) 'transformation design', (B) 'vanishing design' and (C) 'protan/deutan classification'. Please note that these may not be reproduced accurately as the printed colour and the viewing illuminant will be different.

The first 25 plates consist of one introductory plate that demonstrates the visual task, 20 plates for screening and four plates for protan/deutan classification. The transformation, vanishing and hidden digit plates are all intended for screening. The transformation plates contain two figures embedded in the background, one that is read by the normal observer and the other has the appropriate chromatic and lightness contrast that is read by the colour defective observer (Fig. 2.1A). The vanishing design contains a numeral, which is correctly read by normal observers and not seen by colour defective observers (i.e. the 'vanishing' figure is pseudoisochromatic to the colour defective) (Fig. 2.1B). The diagnostic or classification plate is a more sophisticated version of the vanishing design plate that allows for differentiation of a protan from a deutan observer (Fig. 2.1C). Previous studies have shown that the most efficient plates for screening are only the 16 transformation and vanishing plates (Birch 1997c; Cole 1963). The hidden digit designs contain numerals which are camouflaged for normal subjects and not seen, and designed so that red/green colour vision deficient subjects can see them.

The examiner turns the pages and each plate is viewed for approximately four seconds. The observer is told to read the numbers that he/she sees. The instructions given to the observer emphasise the fact that sometimes there is no number on a page.

The Ishihara test is not designed for examining tritan deficiency and so will not identify yellow/blue deficiencies, which are usually predominant in acquired deficiency and any red-green deficiency identified will not necessarily be able to be defined as congenital or acquired without further examination on other colour vision tests. The severity of the defect can be estimated roughly according to the number of errors made. If some of the transformation plates or vanishing plates are read correctly, the defect is likely to be slight. This test does not distinguish dichromats from severe anomalous trichromats.

2.2.2.2 American Optical - Hardy, Rand and Rittler test

The American Optical - Hardy, Rand and Rittler pseudoisochromatic plates (AO-HRR) (Hardy et al. 1954) is still one of the most popular plate tests despite being now out of print. It has both screening and diagnostic plates for tritan defects, and diagnostic and grading plates for protan and deutan defects.

There are four introductory plates used to describe the test, then four protan/deutan and two tritan screening plates, and these are followed by diagnostic and grading plates for protan, deutan and tritan defects. 24 plates are used in total. The plates have vanishing designs containing geometric shapes (circle, cross and triangle) that are printed in neutral colours on a background matrix of grey dots (Fig. 2.2). The saturation of the neutral colours (see Section 1.6.2.3) increases in successive plates to produce designs with progressively larger colour difference steps (Hardy et al. 1954) so identifying different levels of deficiency. This test is often used complementary to the Ishihara plates.

The AO HRR plates are given in reverse order, after showing the introductory plates, so giving the advantage that the test moves from an easy to a more difficult level.

The value of the AO-HRR is in classifying protan and deutan defects, grading the severity of red/green colour deficiency and identifying moderate tritans. The classification of protan/deutan defects is not 'straight-forward' as it is based on

the number of protan/deutan designs failed and, due to the relatively small numbers of plates, people often fail an equal number of both designs.

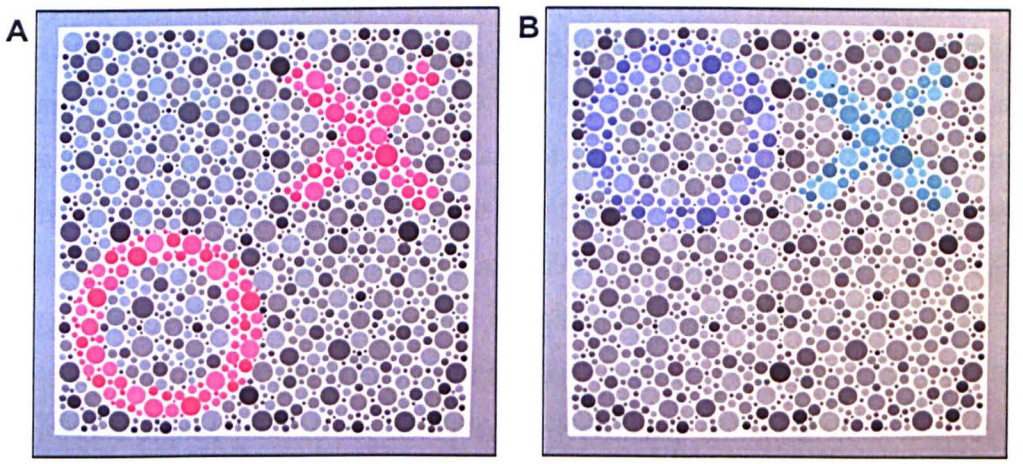


Figure 2.2: The American - Optical Hardy, Rand, Rittler pseudosichromatic plate test. (A) The intended design of this plate was to identify severe protan and deutan deficiency depending on which geometrical shape is seen. (B) Example of a plate intended to identify tritan deficiency. Again note that these may not be reproduced accurately as the printed colour can be different.

2.2.3 Hue discrimination or arrangement tests

In hue discrimination tests, the observer is required to identify colour differences or arrange colour samples according to hue, lightness, and saturation. Arrangement tests were originally designed to identify people with significant colour deficiency who are likely to experience practical difficulties in specific occupations.

2.2.3.1 The Farnsworth-Munsell 100 hue test

The Farnsworth-Munsell 100 hue test (F-M 100 hue test) (Farnsworth 1943; Farnsworth 1947) is one of the most widely used hue discrimination tests. It is not reliable for screening and only people with moderate or severe colour deficiency are identified (Birch 1989). The purpose of the test is to measure hue

discrimination ability for vocational guidance. The test consists of 85 hues, after the test was modified (Farnsworth 1957), and is arranged in four boxes of 21 or 22 colours each. Each box is arranged in turn and the subject's task is to put the colours back in each box in what is perceived to be a natural colour order, starting either from a single reference point or arranging the colours between two reference caps. Caps can be moved freely during the test and are numbered on the back to record the resultant order. An error score can be calculated and this demonstrates the subject's ability to make careful observations. There are no confusion colours from opposite sides of the hue circle in a box and there are less than three colour difference steps between adjacent hues, making the test difficult even for normal trichromats to complete without error. The correct arrangement of the colour samples not only requires normal colour vision but also good chromatic discrimination sensitivity. It has been shown that people can improve their score on re-test, through learning effects. This is an important consideration when the F-M 100 hue test is to be used to monitor changes in severity of acquired colour deficiency. Characteristic F-M 100 hue plots in congenital protan, deutan and tritan defects show concentrations of errors in two positions that are nearly opposite along isochromatic lines tangential to the hue circle. Large differences in the range of error scores are obtained and an error score of more than 100 coupled with an axis of confusion indicates significant colour deficiency (Birch 1989).

2.2.3.2 The Farnsworth-Munsell D-15 test

The Farnsworth-Munsell D-15 (D15) was originally designed for vocational guidance to select recruits with adequate discrimination for work in the electronics industry (Farnsworth 1947). The aim of the test is to separate (dichotomise) those people with moderate and severe chromatic discrimination loss, who fail the test, from people with normal colour vision and slight colour deficiency, who pass the test. In addition, the test indicates yellow/blue discrimination loss, detects achromatopsia (see Section 1.6.3.1), and is useful in the evaluation of acquired colour vision defects (Verriest 1963).

The D15 panel contains 15 movable colour samples. One reference colour sample subtends a visual angle of 1.5° at 0.5m when shown at a distance comfortable for manipulation. The samples are chosen to represent approximately equal hue steps for the natural colour circle. The test contains colours selected from an incomplete Munsell hue circle. Isochromatic colour confusions are demonstrated when colours from opposite sides of the hue circle are arranged side by side in the subject's arrangement.

All the colour caps, except for the reference colours, are removed from the box and mingled on the table. The subject is instructed to place the caps back into the box arranging them "in a natural order according to colour" (caps are numbered on the reverse). The test is performed under appropriate illumination (see Section 2.3.1.3 below). The majority of individuals with normal colour vision can complete the box in one minute. Usually, as much time as necessary is allowed to complete the task. The order of the caps is transferred to the results diagram that shows the correct cap positions forming a circle starting with the reference cap (Fig. 2.3).

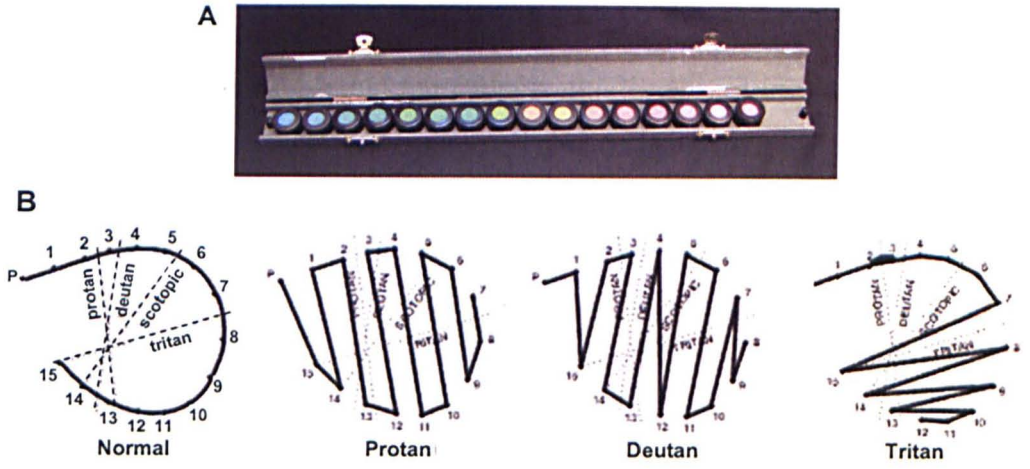


Figure 2.3: (A) Farnsworth D15 test showing a typical sequence ordered by a protanomalous observer. (B) Results plotted on the results form showing typical results for a normal, protan, deutan and tritan observers.

Major errors give rise to lines crossing the circle, e.g. cap 3 and 12 are placed next to each other. These types of errors fall on consistent axes corresponding to protan, deutan, or tritan defects. The axes are indicated in Fig. 2.3B.

2.2.3.3 The City University test (2nd Edition)

The City University test (CU) (Fletcher 1972) consists of a series of ten plates. On each plate five circles mounted on black matt background are displayed: a central and four peripheral colours of equal size. Subjects are asked to select which of the four peripheral colours is most similar to the central colour. The colours employed are all selected from the Munsell series. Three peripheral colours are selected from the opposite side of the hue circle used in the D15 panel sequence and represent typical isochromatic confusions with the central colour, for each type of defect. The fourth colour is an adjacent colour in the D15 sequence and is designated as the normal choice (Fig. 2.4). Since it is based on the D15 panel, the CU test is not designed for screening; approximately 20% of colour defectives pass the test (Birch 1997b) and the protan/deutan classification is not always clear. Most moderate and severe defects are diagnosed correctly and the number of errors is related to the severity of colour deficiency.

The CU test is useful for identifying significant colour deficiency when a format other than the D15 is required.

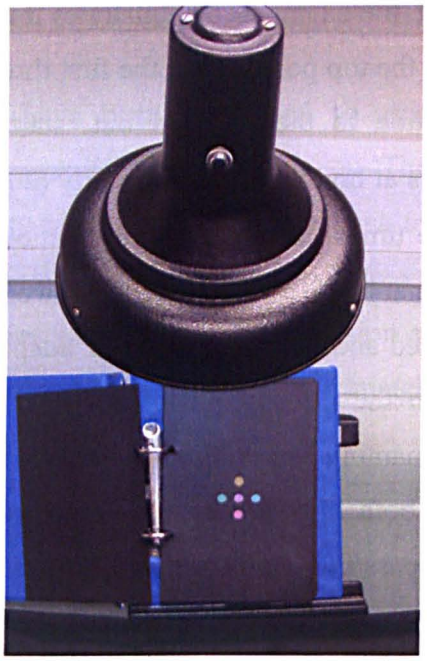


Figure 2.4: Photograph of the City University test illuminated with the Macbeth easel lamp.

2.2.4 Lantern tests

Lantern tests are vocational tests for aviation and the navy that employ colour naming. Some lanterns display single colours and others show colours in pairs. The aim is to determine if signal lights can be identified correctly. The Holmes-Wright lantern, the Farnsworth lantern and the Beyne lantern are examples of lanterns in use that are currently approved by the CIE. All three lanterns follow the same general principle and only the Holmes-Wright lantern will be described in detail.

The Holmes-Wright (type A) lantern (Holmes and Wright 1982) is an aviation lantern showing two vertical colours (Fig. 2.5). Two reds, two greens and one white are used, which have x,y chromaticity co-ordinates within the internationally agreed specifications for signal lights (CIE 2001a). Nine pairs of the colours are shown representing all the possible colour combinations. The lanterns are viewed at 6 m (20 ft).

Before beginning the test, the examiner demonstrates the colours by showing red, green and white lights in the top positions of the first three pairs and naming them correctly. The examination is then carried out, starting in fairly dim room illumination, with colours at high luminous intensity (200 μcd). The nine pairs of colours are shown three times, making 27 presentations of about five seconds each. The subject has to name the top and then the lower colour shown each time. The room is then darkened and the subject is dark adapted for about 15 minutes; the test is repeated with the three sequences of nine colour pairs again. The examiner notes each misnaming of a colour. The results are sorted into number of types of misnaming for both light levels, e.g. 3 greens called 'white', 2 whites called 'green', etc. The pass/fail limits adopted by the Joint Aviation Authority (JAA) consist firstly in showing one run and if all nine pairs of lights are named correctly, no further runs are carried out and the subject passes the test. If there are any errors, two more runs are carried out and if the subject makes no errors in these 18 pairs the test is passed. However, if any mistakes are made then the subject is dark adapted for 15 minutes and one more run of nine pairs is undertaken in the dark. If the subject makes no errors on this final run, the results are taken as a pass. If at any stage the subject calls red 'green' or green 'red', then the test is discontinued and the test is failed.

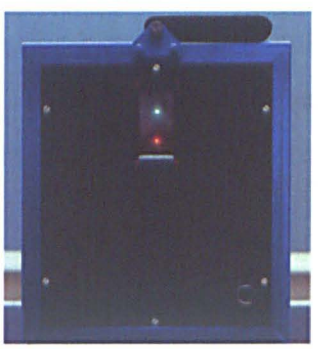


Figure 2.5: Photograph of the Holmes-Wright lantern used in aviation.

2.2.5 Anomaloscopes

Spectral anomaloscopes have been developed to evaluate specialised colour matches, and diagnose congenital colour deficiency. The use of colour matching to establish an abnormality of colour vision dates back to Lord Rayleigh's (1881) discovery that a colour match of a yellow test light to a mixture of red and green primary lights distinguishes variations of colour vision and thus allows diagnosis of congenital red/green colour defects. This colour match can be represented as a linear equation, for example:

$$a_Y Y \equiv a_R R + a_G G \qquad (2.1)$$

where $\equiv$ reads matches, and a_Y, a_R and a_G represent the amount of the primary lights employed in the match. There are two variables: the intensity of the yellow and the relative mixtures of the red and green lights. Ideally the S-cones and rods are excluded so that the match is determined solely by the relative absorptions in the M- and L-cones. This is likely to be so because of the choice of wavelengths employed in the Nagel anomaloscope and the restriction of the stimulus to the central foveal region.

Subsequently, blue-green equations, similar to the Rayleigh red-green equation were derived to test for yellow/blue colour defects. This was the Engelking-Trendelenburg equation, optimised for the tritanomalous match by Moreland and Kerr (1978). There are several problems that limit the diagnostic ability of yellow/blue defects using colour matching. The equations are also affected by individual variations in lens and, to a lesser extent by the macular pigment. One result of this variability is that Engelking-Trendelenburg equation shows a wide distribution of colour matches for normal colour vision observers, thus decreasing its reliability as a test for yellow/blue defects.

2.2.5.1 Nagel anomaloscope

The Nagel anomaloscope is the standard reference test for identifying and diagnosing red/green colour deficiency despite being currently out of production.

The efficiency of other red/green colour vision tests is often compared with the results obtained on the anomaloscope.

The Maxwellian view instrument projects an image of a circular split field onto the retina by first focusing the light into a small portion of the eye's pupil (Fig. 2.6). Minor focusing adjustments can be made with a telescope tube. The field size ranges from 1.8° to 2°. The top half of this circle is illuminated by a mixture of spectral red and green wavelengths; 670 and 546 nm, and the lower half illuminated by spectral yellow, 589 nm. Two control wheels are used, one to alter the red-green colour mixture ratio of the top field, and the other to alter the luminance of the yellow lower field. The red-green mixture wheel can be altered from 0 (only spectral green present) to a value of 73 (only spectral red present). The luminance of the mixture field is kept constant for any red-green mixture ratio. There are minor variations in the peak wavelengths from one instrument to another, thus before it can be used, the normal matching range has to be established for each instrument by examining a large number of normal subjects. Normal trichromats make a precise colour match within a small range of red-green mixture ratios. This means that mixture ratios obtained on different instruments are not always equivalent.

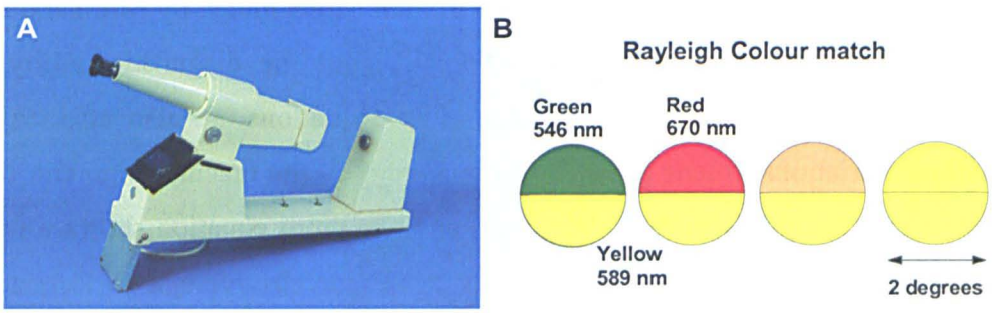


Figure 2.6: (A) Photograph of the Nagel anomaloscope (Model I, Schmidt and Haensch, Germany). (B) Illustration of the Nagel anomaloscope split field. The percentage mixture of red to green in the top half and the luminance of the yellow bottom field can be changed until a match of the two fields can be perceived.

The examiner instructs the observer to make several exact colour and luminance matches. The matching range is recorded from the matching limits on the red-green mixture scale. The range on the Nagel is defined as: $(x_2 - x_1) + 1$ where x_2, x_1 are the upper and lower limits of the red-green match, respectively. The midpoint is the middle value between the matching limits, i.e. a typical match would be 39-41 giving a range of 3 and midpoint of 40. The Rayleigh match for dichromats, like all colour vision tests, exploits the dichromat's colour confusions. As they have only one pigment, or only one functionally distinct pigment in the spectral range provided by the instrument, they have one degree of freedom in the Rayleigh equation and are able to fully match the spectral yellow primary to any mixture of the spectral red and green primary lights by merely adjusting the intensity of the yellow, independent from the red-to-green ratio. Their matches will cover the whole range including both the red and green primaries. Although the deuteranope will display normal or near-normal relations in the luminosity of the matches, the protanope will display a luminosity loss for the red primary, requiring less yellow light to match it. The protanopic luminosity loss for long wavelengths is expressed on the Nagel in the different luminance settings of the 589 nm light used in the matches (Fig. 2.7) - the 589 nm light is set at a relatively dim level in the match to 670 nm and at a relatively high level in the match at 545 nm. The approximately normal deuteranopic luminosity is represented by approximately equal luminance settings of the 589 nm test light to either primary or all mixtures (Fig. 2.7). The classification on the Nagel anomaloscope can be made according to the extent and location of the matching range (Franceschetti 1928).

The regression line fitted to the matches of either the protanope or the deuteranope, its slope and intercept with the yellow intensity axis will depend on the λ_{max} of the underlying photopigment. A protanope has a steep slope whilst the deuteranope has a flat slope (Fig. 2.7).

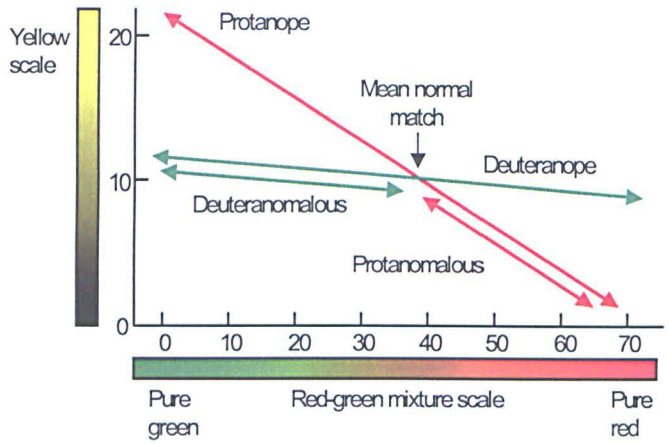


Figure 2.7: Schematic of the type of colour matches for protan and deutan colour vision deficient observers obtained on a Nagel anomaloscope. Protanopes and deuteranopes are able to match the yellow with all red-green mixture ratios by adjusting the luminance of the yellow comparison field. Protanomalous trichromats obtain matching ranges with an excess of red and deuteranomalous trichromats obtain matches with an excess of green in the matching field. The normal match is usually not accepted in anomalous trichromatism.

In summary, the colour match on the Nagel anomaloscope gives accurate information about the type and severity of the colour vision deficiency. The important parameters are the end points of the red-green mixture scale that determines the extent and midpoint of the matching range and the luminance value needed for the match. Several colour matches are needed to establish the full range of primary mixtures that are acceptable and determine accurately the matching range (Schmidt 1955).

2.3 Evaluation of colour vision tests

The purpose of this study is to compare and evaluate several colour vision tests, such as the Ishihara plates (Kanehara Shuppen Co. Ltd., Tokyo, Japan), the AO-HRR (New York, USA), the D15 (Richmond Products, USA) and the CU (Keeler Ltd., UK), with respect to a reference test, the Nagel anomaloscope. The aim was also to examine the consistency of results on the various tests in relation to the pass/fail criteria set for the aviation environment.

2.3.1 Material and methods

2.3.1.1 Subjects

Normal and colour deficient subjects were recruited by advertisement or by random selection from the university environment and by referral from other colour vision clinics. There were 472 subjects in total, 345 men and 127 women who were tested and whose results were compared. The study protocol was approved by the Research and Ethical Committee at City University and each subject provided informed consent before participating. Each subject was asked before starting the tests whether they suffered from any health problems. Positive answers meant that the data were treated separately from the rest. Table 2.2 gives the number of subjects that carried out the different tests. Not all subjects took all tests, because of time constraints and technical difficulties with the Nagel anomaloscope towards the end of the study.

Tests	No. of subjects
CAD test	472
Nagel Anomaloscope	231
Ishihara (25 plates)	311
AO-HRR	242
D15	240
City University (2nd Ed.)	226

Table 2.2: Number of subjects that carried out the different colour vision tests described in the text.

The age range of the subjects tested was from 7 to 75 years old, with a median of 28 years old. The distribution of subject ages is shown in Fig. 2.8 below.

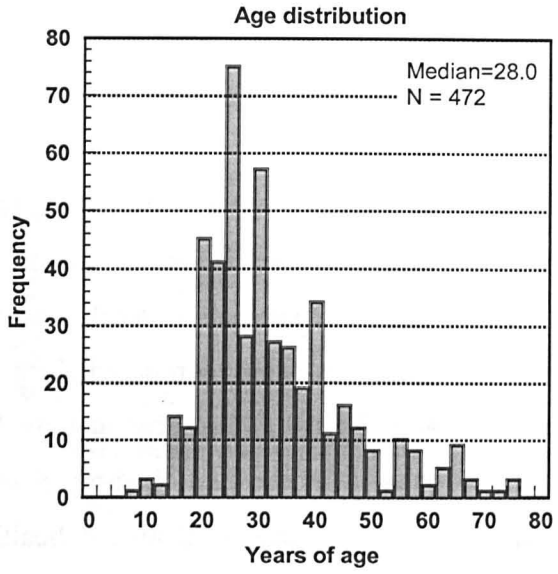


Figure 2.8: Frequency distribution of subjects tested in this study. The ages range from 7 to 75 years old with a mean of 31, standard deviation of 13 and a median of 28 years old.

2.3.1.2 Administration of table based colour vision tests

Examination procedure

It is important that the same procedure and conditions are used for administering each test. This then allows comparisons to be made between data collected on different occasions or by different examiners. Instructions to the subjects and the level and type of illumination remained unchanged throughout the study. The results from the battery of tests were collected together to provide an accurate assessment of the subject's colour vision.

Illumination

Chromatic differences can change with the spectral content of the illuminant. The majority of the pigment tests described above were designed and standardised for natural daylight or for the CIE standard illuminant C (see Section 1.5.2). 'Natural daylight', i.e. northern sky light in the northern hemisphere, may be used for colour vision tests, however it varies with the time of year and day. A more

practical form of illumination is the Macbeth Easel lamp (Kollmorgan Corp., Waltham, Massachusetts) especially designed for plate tests (see Fig. 2.4). The Macbeth lamp consists of 100-W tungsten bulb fitted with a hemispherical blue filter that converts the spectral composition to standard source C, providing 350-400 lux on the surface of the plate positioned on a specially designed tray, at 45°, below the lamp.

Colours specified within an isochromatic zone under source C will occupy different positions in the CIE chromaticity diagram for different illuminants. Previous studies have shown that unfiltered tungsten illuminants with colour temperatures lower than source C results in deuteranomalous trichromats making fewer errors on pseudoisochromatic plates, such as the Ishihara and the AO-HRR tests (Hovis and Neumann 1995).

Pass/fail criteria

The pass/fail criteria for the different tests will affect the efficiency of a given colour vision test. Here, the criterion for passing the Ishihara test was set at making no errors (JAA protocol), accepting no *potential* misreading, in any plate from 1 to 25. The 'font' of the numerals produced often leads observers with normal colour vision to 'fill in' partial loops in the design. As a result, a '5' may be interpreted as '6' or a '9' as '8'. In order to prevent any possible misinterpretation, no errors were permitted. For a list of misreadings used by London Underground Occupation Health refer to Appendix D). However, in the literature there are studies that recommend allowing up to three misreadings on certain plates because according to (Birch 1997c) about 40% of normal observers make at least one misreading and these are separate from clear errors made on other plates. Further, Birch (1997c) recommends a criterion for the pass/fail limit on the Ishihara test, as a very efficient screening test (not intended for occupational colour vision screening), to consider only the 'transformation' and 'vanishing' plates (i.e. the first 16 plates) and allowing up to a maximum of three misreadings but no errors (in a manner indicating red/green colour deficiency). Colour deficient people record several errors and due to the large number of

screening plates, and the different types of design, there is little uncertainty in the overall result.

Although previous studies have recommended allowing one single error on any of the four screening plates (Birch 1997a), in this study, a pass on the AO-HRR plates was set at making no errors in any of the 20 plates.

Similarly the criterion for passing the CU test was also set at zero errors in any of the 10 plates. For the D15 no crossings between colour caps was allowed, however up to three transpositions of adjacent caps was permitted (see Fig. 2.3).

2.3.1.3 Setting the norm and calibration of the Nagel anomaloscope

The norm of the reference test has to be established in order to make comparisons between different colour vision tests. The definition of the 'normal observer' on the Nagel anomaloscope will allow distinguishing between normal and colour deficient vision and thus evaluate another colour vision test with its specific pass/fail limit.

To define normal trichromacy and diagnose colour deficiency, there are characteristic colour matches for normal and congenital red/green colour defects. Fig. 2.9 shows the anomaloscope matches performed on 231 subjects. The distribution of range widths and midmatching points is very variable particularly amongst anomalous trichromats. Typical mean yellow luminance values (see Fig. 2.7) found in protanopia were 2 ± 0.6 and in deuteranopia 15 ± 2.5. Within the protan defects, three subtypes can be recognized: simple protanomalous trichromats (abbreviated PA) make a unique match that contains a higher proportion of red primary compared to the normal match. The extreme protanomalous trichromat (abbreviated XPA) has a wide matching range that includes the simple protanomalous match and the normal match and may extend to one of the primaries. The protanope (abbreviated P) as a dichromat matches using different values of luminance of the test light to both primaries (high and low luminances for red and green primaries respectively, see Section 2.2.5). For

the deutan deficiencies, the simple deuteranomalous trichromat (abbreviated DA) makes a unique match that contains a higher proportion of green primary than the normal match. The extreme deuteranomalous trichromat (abbreviated XDA) has a wide range that includes the simple deuteranomalous match and the normal match and may extend to one of the primaries. The deuteranope (abbreviated D), as a dichromat matches the yellow test light to the whole matching scale with equal values of luminance.

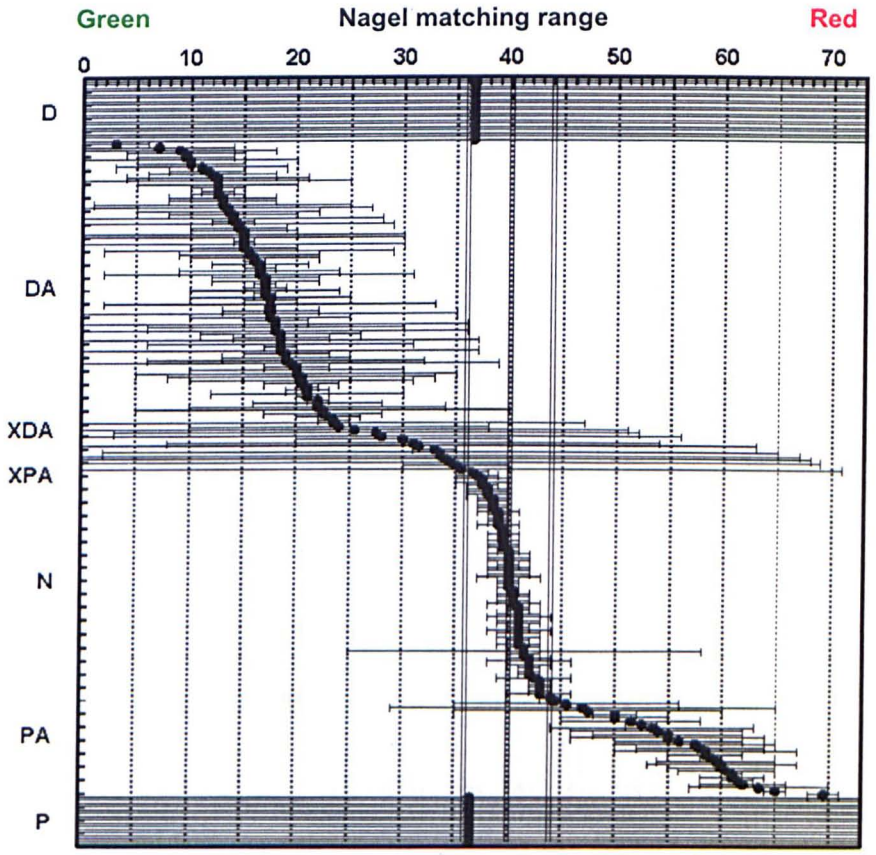


Figure 2.9: Matching ranges and midpoints for 231 subjects measured on the Nagel anomaloscope. D (deuteranopes), XDA (extreme deuteranomalous), DA (simple deuteranomalous), N (Normal), PA (simple protanomalous), XPA (extreme protanomalous), P (protanopes). XPA and XDA are distinguished due to differences in the yellow setting of the matching range.

The midmatching point and the matching range both have theoretical significance. When an anomalous trichromat accepts a narrow match, the midpoint is determined by the absorption spectra of the visual photopigments and the lens transmission function. The matching range is an indication to the observer's chromatic discrimination ability, i.e. an important measurement in the practical context. On average, the matching range of an anomalous trichromat is wider than that of the normal trichromat, however occasionally anomalous trichromats have very narrow matching ranges. Fig. 2.10 shows a distribution of the midmatching points for 197 subjects, including normals and anomalous trichromats (dichromats have been excluded). From this distribution three different groups of midpoint matches can be identified, those with midpoint values less than 30 on the red-green mixture scale, between 30 and 48 and those having midpoints greater than 48.

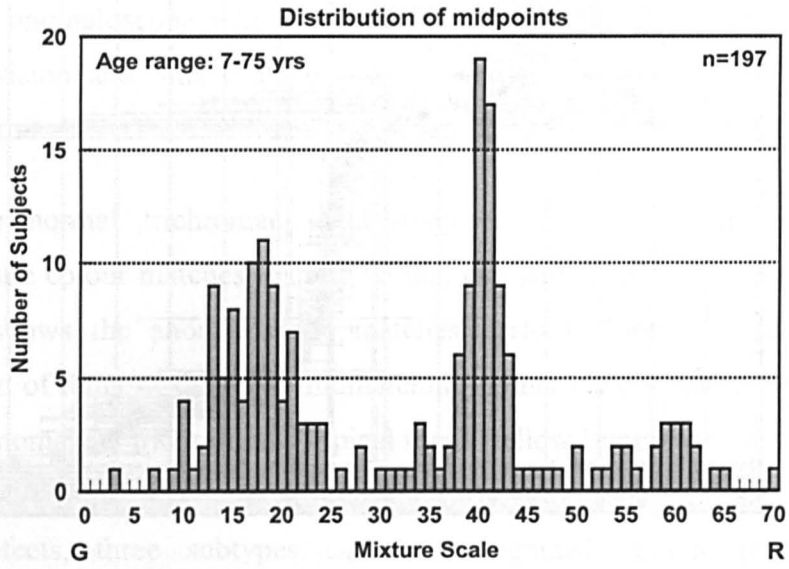


Figure 2.10: Distribution of the midpoint values of the red-green mixture range for 197 subjects, ranging from 7 to 75 years of age. The distribution of midmatching points is symmetrical for normals with a mean value near 40 scale units. The midmatching points for anomalous trichromats are located on both sides of the normal midpoint. Dichromats have been excluded from this plot as their midpoint will always be situated at 36.5 and will obscure the data.

Using this preliminary identification of three groups, Fig. 2.11 shows the variability in the widths of ranges. The matching ranges for both those groups with midpoints less than 30 and those above 48 are very variable, extending from extremely narrow ranges of 2 units to up to 72 units. Those subjects with midpoints 30 to 48 (population of normal observers) have relatively less variable widths of ranges and their midpoints show a symmetrical peaked distribution (see Fig. 2.10) and such a distribution can be characterized by its mean and standard deviation. Different normal trichromats will use slightly different amounts of the primaries to match various hues, but it is the general similarities rather than the comparatively small differences among normal observers that allow us to classify observers as either having normal or deficient colour vision.

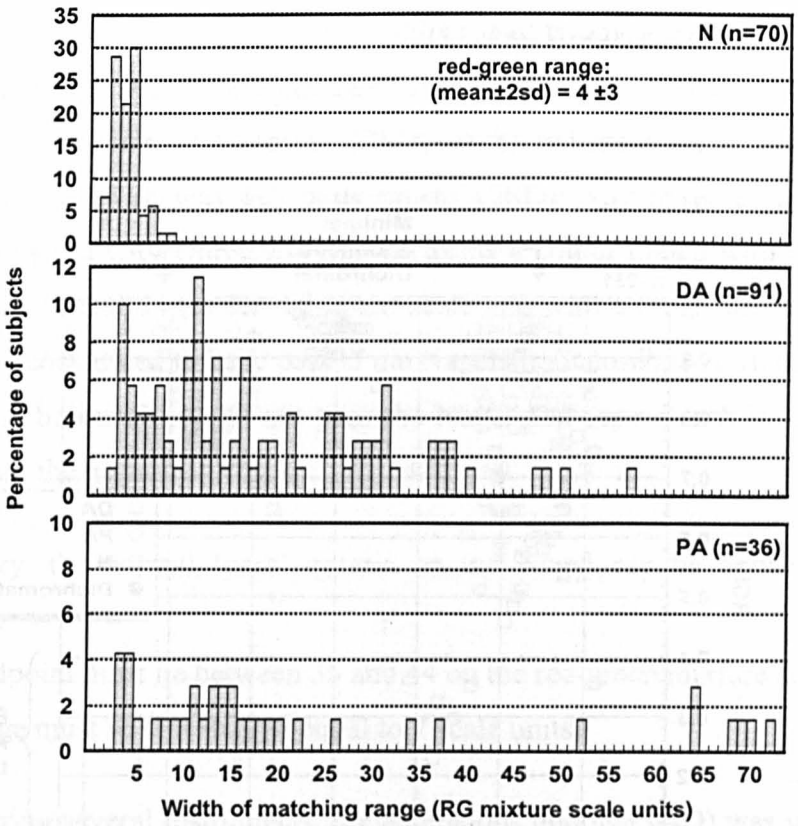


Figure 2.11: Histograms showing the percentage distribution of matching ranges among the normal (N), deuteranomalous (DA), and protanomalous (PA) subjects.

Previous studies have shown the lack of correlation in the Nagel anomaloscope between the matching range and the midpoint for normal and deuteranomalous trichromats, even though Rayleigh equations for protanomalous trichromats have shown some dependence between range extent and midpoint (Pokorny & Smith 1977). A parameter relating to the matching range, which provides indication of the observer's chromatic discriminative ability, is introduced in this study, the red-green discrimination index (RGI), which is defined as follows:

$$1-\frac{range}{74}=RGI \qquad (2.2)$$

Although midpoint and range have been shown to be only weakly correlated (Hurvich 1972), a scatter plot of Nagel midpoints on the red-green scale versus RGI allows to separate a clear cluster of subjects, with midpoints between 36 and 44 on the red-green mixture scale units, that are most likely to form our normal group on the anomaloscope (Fig. 2.12).

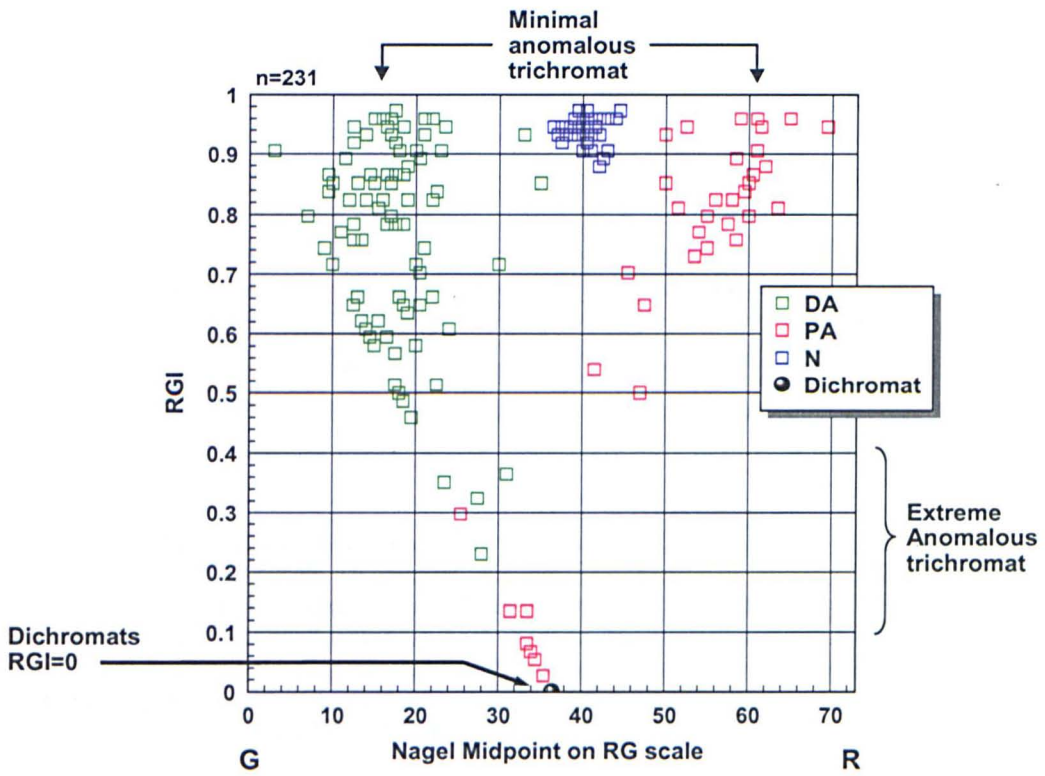


Figure 2.12: Scatter plot of matching midpoints versus RGI for all observers (n=231). See next page for further details.

Figure 2.12 continues: 70 subjects formed a cluster that is separated from all other subjects by having a midpoint around 40 and a high RGI. The abbreviations in the legend refer to deuteranomalous (DA), protanomalous (PA), and normal trichromat (N) subjects. The value of the luminance setting on the yellow scale provides additional information to separate deutan and protan colour deficient observers. In Appendix F, a 3D plot showing the yellow setting of the red-green match for each observer has been included.

The match midpoint and matching range parameters for 70 subjects formed a distinct cluster that was separated clearly from all others. Fig. 2.12 shows that the majority of observers with red-green midpoints from 30 to 48 also have high RGI values, i.e. narrow matching ranges, labelled 'N' (see also Fig. 2.11). Four subjects are labelled PA despite having midpoints between 30-48 on the red-green scale due to their relatively low RGI. The parameters of the matching range and midpoint measured for the observers (N), shown in blue, were used to calculate the statistical limits that describe the 'normal' trichromat on the Nagel anomaloscope. The mean midpoint value and mean matching range was estimated together with the standard deviation of the means. The mean red-green setting, midpoint, was 40 ±4 (mean ±2SD) on the red-green scale; and the mean red-green range width was 4±3 units (mean ±2SD). Any observer when tested with the Nagel at City University that performs a colour match with a midpoint lying within 36 and 44 on the red-green scale and with a range width of 1 to 7, will thus be considered to have passed the Nagel anomaloscope. Two subjects are labelled DA because they do not pass the Nagel and their match is sufficiently distinct from the normal cluster.

In summary, the normal 'pass' criteria on the Nagel can be defined by two conditions:
(i) The midpoint must lie between 36 and 44 on the red-green mixture scale and,
(ii) the range must be less than or equal to 7 scale units.

For comparing several instruments, the anomalous quotient (AQ) was introduced by Trendelenburg (1929) for compensating for minor changes in line voltage and bulb aging. Small fluctuations of the order of 0.2 Nagel units per °C in colour matches been shown to occur due to seasonal variations that cause a change in refractive index (Jordan and Mollon 1993b). AQ involves calculation of the

numerical value of an individual mean match relative to the mean match value obtained for a large group of normal subjects. It can be calculated from the midpoint of a matching range. For example; if the mean normal match for an instrument is 40 on a red-green mixture scale (pure green at 0 and pure red at 73) and the mean for an individual observer is 38, then his AQ is:

$$AQ = \frac{73-38}{38} \Big/ \frac{73-40}{40} = 1.12 \qquad (2.3)$$

AQs are greater than unity when there is an excess of green in the match, and are less than unity when there is an excess of red in the match compared with the mean value. The AQ is completely equivalent to the midpoint and carries no information regarding the extent of the matching range, so therefore is inadequate for classifying colour deficiency. The AQ is useful for comparing different instruments with different red-green mixtures, and for defining the normal range when the light source has been replaced.

The definition of the Nagel normal observer agrees with typical parameters described in the literature. The AQs for the normal observer would range from between 0.80 and 1.25 on the instrument used in this study. Similarly, Helve (1972) found for 186 normal observers a range of AQs between 0.80 and 1.20. In the French and Belgium armies a subject is considered to have a normal colour vision when he has a mean range from 4 to 8 units and his matching point has an AQ between 0.63 and 1.3, which corresponds to midpoints ranging from 35 to 48 on the red-green scale (written communication from Frans Larminier). The JAA considers a pass on the Nagel anomaloscope to be a range of 4 units or less (JAA 2002).

2.3.1.4 Statistical analysis

Colour vision tests serve different functions and several tests may be needed for a full colour vision assessment. The tests selected may depend on the priorities of the examination; i.e. for screening, grading or classifying. Conventionally,

methods of examination involve several different visual tasks such as verbal identification of a figure or colour matching. The efficiency of most tests in current use has been established in clinical trials and with the use of a Nagel anomaloscope as the standard reference or 'gold standard' for determining normal or abnormal red-green colour vision and diagnosing the exact type of red/green colour deficiency. The efficiency of a colour vision test can be investigated with reference to the anomaloscope result, by calculating the proportions of subjects with normal and abnormal colour vision who are likewise 'diagnosed' by the two tests. The terms 'sensitivity' and 'specificity' are used to indicate screening efficiency. These parameters for describing diagnostic tests are formally defined with relevance to colour vision as follows (Altman 1991):

- *Sensitivity* is the proportion of colour defectives that are correctly identified by the test.
- *Specificity* is the proportion of observers with normal colour vision that are correctly identified by the test.

An efficient colour vision screening test should have a specificity of 100%, which suggests that no people with normal colour vision are incorrectly identified as colour deficient (false positives). Sensitivity should be preferably high; a value of 90% is equivalent to 10% of colour deficient people incorrectly identified as normal (false negatives). The necessary level of sensitivity and specificity depends on the clinical screening situations, for example in cancer screening the sensitivity should be high ~100% so that no 'abnormals' are missed and further confirmatory tests are often performed to exclude false positives. In colour vision screening a small percentage of false negative results is acceptable as long as these usually represent cases of very slight, or minimal, colour deficiency that are not involved in occupational colour critical tasks.

Once the result of the colour vision test is known, it is also useful to know what the expected probability of normal or abnormal outcome. The effect of prevalence of the abnormality is a factor inherent in the above. The observed prevalence of colour vision deficiency depends on the environment in which the testing is carried out and the motivation behind the testing. For this study, there are two

most typical situations that may be relevant: (i) screening of subjects from the general population with a low prior probability of a colour vision deficiency and, (ii) attendance in a colour vision clinic (i.e. patients who are referred as a result of previously failing a colour vision assessment). A formal definition describing this in relation to colour vision follows (Altman 1991):

- *Positive predictive value (PPV):* is the proportion of observers with a diagnosis of deficient colour vision that is correctly diagnosed.
- *Negative predictive value (NPV):* is the proportion of observers with a diagnosis of normal colour vision that is correctly diagnosed.

Mathematical definition of these parameters can be found in Appendix C. These values give a direct assessment of the usefulness of the test in practice.

2.3.2 Results

2.3.2.1 Efficiency of principal colour vision tests

If the definition of the "normal observer" on the Nagel anomaloscope is accepted then comparisons with the different colour vision tests can be carried out to evaluate the agreement with the Nagel. These results are summarised using square frequency tables.

Table 2.3A shows the comparison of the Ishihara plate-test (accepting no errors) against the Nagel. The results show 100% sensitivity: correct identification of all colour deficient observers; and 87.5% specificity: incorrect identification of some colour normal as defective. This could suggest that the Ishihara is a more stringent colour vision test in comparison to the Nagel anomaloscope. The data show three subjects that fail the Ishihara but fulfil the normal criteria on the Nagel; two of the subjects fail by making one error and one misreading each, and the third subject fails by making two misreadings but no errors. The Nagel results for the first two subjects are well within the normal observer limits, both with ranges of 3 units and midpoints at 40 and 41, while the third subject has a Nagel

range of 7 units and midpoint of 40 (also within the normal observer limits). If three misreadings were permitted in the first 25 plates, the specificity would increase to 91.7% (see Appendix A).

Table 2.3B shows results for the AO-HRR test in comparison to the Nagel. In this case the specificity is 100%, unlike the Ishihara test, which indicates that all normals are correctly identified. The pass/fail criterion for the AO-HRR is based on no errors in any of the 20 plates which results in a sensitivity of 94.2%. The nine subjects that fail the Nagel but pass the AO-HRR were all deuteranomalous with midpoints below 20 on the red-green mixture scale and ranges varying from 3 to 11 units. If one error is permitted in any of the screening plates (as suggested in a previous study), the sensitivity would drop to 84.4%, so in this study the AO-HRR criterion of making no errors provides the highest screening efficiency.

Table 2.3C shows results for the D15 in comparison to the Nagel. For this test the specificity is 100% indicating that all normals are correctly identified. The sensitivity is only 53.5% allowing almost half of the observers with abnormal colour vision to pass this test.

Table 2.3D shows results for the CU test in comparison to the Nagel. For this test the specificity is 100% indicating that all normals are correctly identified. The sensitivity is 60.5% revealing that many observers with abnormal colour vision can pass this test without making any error.

A

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

Ishihara test
Standard version 38-plate
1-25 plates (no errors & no misreadings)

	Pass	Fail
Pass	21	0
Fail	3	156

κ=0.93
Specificity=87.5% (95% CI: 0.67-0.97)
Sensitivity=100% (95% CI: 0.97-1)
180 subjects

B

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

AO-HRR
(no errors)

	Pass	Fail
Pass	24	9
Fail	0	145

κ=0.84
Specificity=100% (95% CI: 0.83-1)
Sensitivity=94.2% (95% CI: 0.89-0.97)
178 subjects

C

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

D15
Max. of 2 adjacent transpositions

	Pass	Fail
Pass	24	72
Fail	0	83

κ =0.40
Specificity=100% (95% CI: 0.83-1)
Sensitivity=53.5% (95% CI: 0.45-0.62)
179 subjects

D

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

CU
(no errors)

	Pass	Fail
Pass	24	58
Fail	0	89

κ =0.45
Specificity=100% (95% CI: 0.83-1)
Sensitivity=60.5% (95% CI: 0.52-0.68)
171 subjects

Table 2.3: Contingency tables comparing the Ishihara plates, AO-HRR, D15 and CU colour vision tests with the Nagel anomaloscope. The number of subjects that carried each pair of tests is shown. The pass/fail criteria for each test are shown and the agreement (κ), specificity and sensitivity are calculated. 95% confidence intervals (CI) for proportions are calculated according to the method described by Newcombe (1998) based on the procedure outlined by Wilson (1927).

Sensitivity and specificity describe the probability that subjects with anomalous colour vision are diagnosed as having deficient colour vision and also the probability that subjects with normal colour vision are diagnosed as having normal colour vision. However in an actual test situation one is faced with just the diagnosis and it would be desirable to know the probability that such a diagnosis based on the result of a particular test is correct. This is given by the PPV and NPV, which are affected by the prevalence of abnormality. Most of the subjects examined in this study were referred to us because they had been previously identified as colour defective by failing one or several colour vision tests. Thus a high number of colour deficient observers are to be expected from this set of subjects. At City University the prevalence of colour vision deficiency was of the order of 87% compared to typical values for a representative cross-section of the general population: 8% of men and 0.4% of women (see Table 2.4).

			Prevalence City University 87%		Prevalence Males 8%		Prevalence Females 0.4%	
Test	**Sens†**	**Spec††**	**PPV[1]**	**NPV[2]**	**PPV[1]**	**NPV[2]**	**PPV[1]**	**NPV[2]**
Ishihara	100	87.5	98	100	41.0	100	3.1	100
AO-HRR	94.2	100	100	72.3	100	99.5	100	100
D15	53.5	100	100	25.0	100	96.1	100	99.8
CU	60.5	100	100	29.3	100	96.7	100	99.8

† Sensitivity
†† Specificity
[1] Positive Predictive Value
[2] Negative Predictive Value

Table 2.4: Summarised results for several colour vision tests with reference to the Nagel anomaloscope. The sensitivity and specificity calculated for each colour vision test was used to find the PPV and NPV values for the situation found at City University and also for the typical prevalence of colour vision deficiency in the average population for both males and females.

Table 2.4 summarises the results obtained from the colour vision tests used in this study with reference to the Nagel anomaloscope. The Ishihara plate test reveals

100% NPV (95% confidence intervals (CI): 0.81-1) which suggests that all people in our sample who passed this test have indeed normal colour vision, i.e. there were no false negatives. However, 2% of those diagnosed as being colour deficient are incorrectly diagnosed (PPV=98%; 95% CI: 0.94-1), i.e. 2% of those who fail the Ishihara test have normal colour vision (according to the Nagel anomaloscope). For an average population sample, using the values for the screening efficiency in this study, it is suggested that the Ishihara test provides excellent screening efficiency, and males and females who pass this test are correctly diagnosed with normal colour vision (NPV=100% for both males and females). Those diagnosed as colour deficient will need further examination with additional tests to confirm the diagnosis (PPV=41.0% for males and PPV=3.1% for females). The AO-HRR, D15 and CU tests all have 100% PPV (95% CI: 0.97-1, 0.95-1, 0.95-1, respectively) in our sample; no one is incorrectly diagnosed as colour defective. In a colour vision clinic setting (such as the City University environment) further tests will have to be used for confirmatory results of normal colour vision due to the relatively low NPV values: AO-HRR 72.3% (95% CI: 0.54-0.86), D15 25.0% (95% CI: 0.17-0.35) and CU 29.3% (95% CI: 0.20-0.41). In the average population, both the PPV and NPV of the AO-HRR plates are very high. For the D15 there are 4% of people diagnosed as being colour vision deficient that may be wrongly diagnosed so again further testing will be necessary for correct identification of those that are supposed to have a colour vision deficiency. Similar results apply to the CU test.

It must be noted here that the number of subjects diagnosed as normal on the Nagel anomaloscope and that carried out the other colour vision tests is relatively small compared to the number of subjects diagnosed as colour deficient by the Nagel. This results in a high prevalence of colour vision deficiency amongst the people that attend City University. And the estimates of 100% specificity for the AO-HRR, D15 and CU may be too high as they are based on a small sample (the lower limit of 95% confidence interval is 0.83 for these tests as shown in Table 2.3).

The agreement, given by κ, is a measure of the agreement between each test and the Nagel anomaloscope that is corrected for chance (see Appendix C). The

colour vision test in highest agreement with the reference test is the Ishihara plates, followed in order by the AO-HRR, CU test and D15 (Table 2.3).

2.3.2.2 Pass/fail comparison

No. of Subjects (total=164)	Colour Vision Test				
	Nagel	Ishihara	AO-HRR	D15	CU
73	F	F	F	F	F
4	F	F	F	F	P
16	F	F	F	P	F
45	F	F	F	P	P
9	F	F	P	P	P
17	P	P	P	P	P

Table 2.5: Comparison of the various colour vision tests for the same 164 subjects that did all five tests. Not all subjects carried out all the colour vision tests.

Table 2.5 shows the number of subjects that pass/fail each test of the battery of different colour vision tests. In this sample of 164 subjects who performed all five tests, we can see that different colour vision tests often give discordant results. For 45% of the subjects at least one of the tests gave a divergent result. All tests gave the same results for 73 subjects who failed all tests and 17 subjects that passed all tests, so for a total of 55% of the subjects there was complete agreement. Most often the disagreement here is between the tests with lower sensitivity (the D15 and CU), which are designed to pick up moderate to severe colour deficiency, and the tests which also diagnose mild deficiency correctly. This is the case for 65 subjects, who passed one or both of the D15 and CU, but failed the other three tests. The AO-HRR has a higher sensitivity than these two

tests, but there are also 9 subjects who pass the AO-HRR (and D15 and CU), but fail the Ishihara plates and the Nagel anomaloscope.

2.3.2.3 Performance comparison against the Nagel anomaloscope

If we look separately at those subjects diagnosed as having deutan deficiency (labelled D and DA) and protan deficiency (labelled P and PA) on Fig. 2.12, we can investigate the number of errors the subjects make on the other tests in comparison to their Nagel range. For larger Nagel ranges (i.e. smaller RGI values), chromatic discrimination is worse (Pokorny et al. 1979).

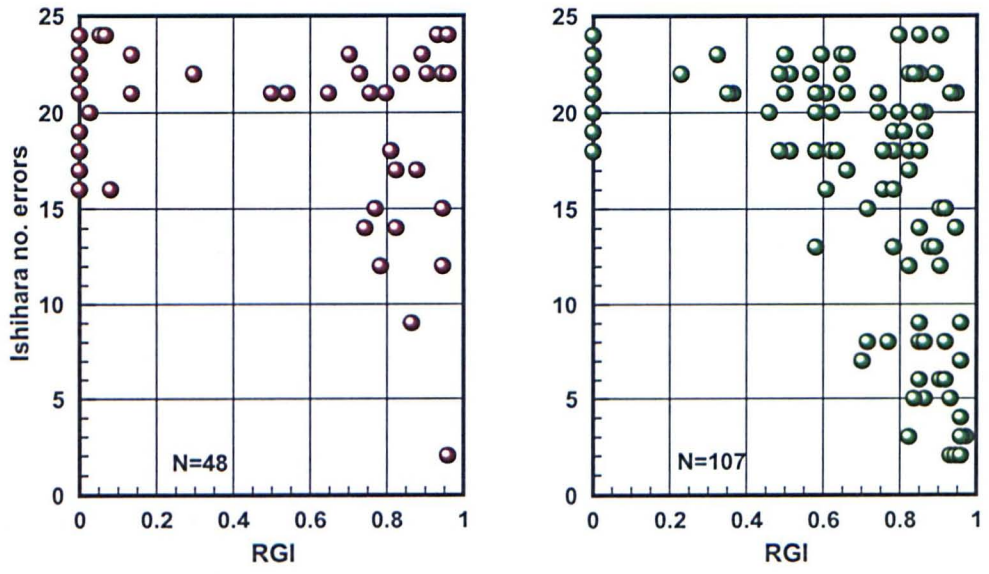


Figure 2.13: Comparison of the number of errors on the Ishihara plate test versus red-green discrimination index (RGI) for protan (left panel) and deutan (right panel) colour vision deficient subjects. The correlation coefficient for best linear fits is (A) 0.11 and (B) 0.22.

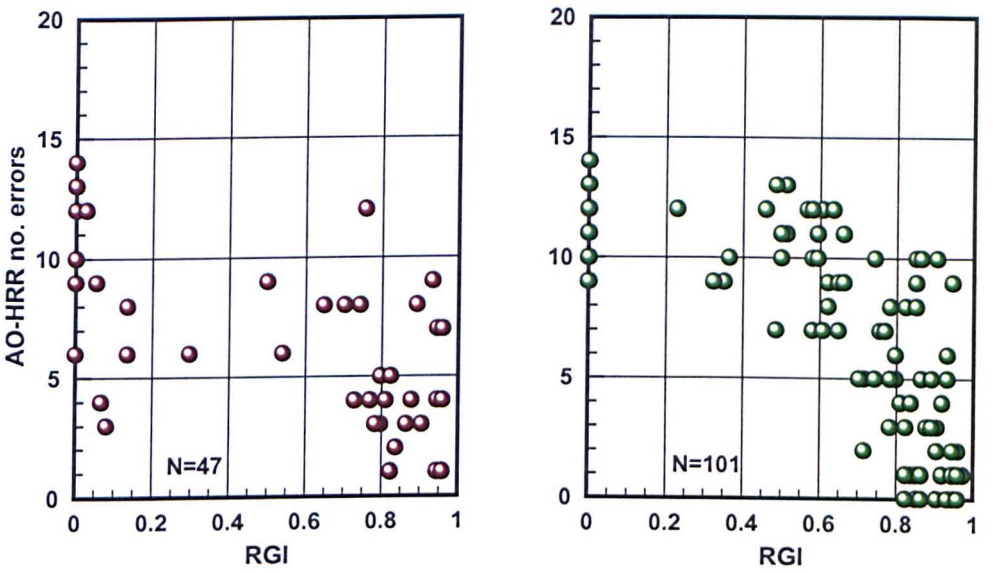


Figure 2.14: Comparison of the number of errors on the AO-HRR plate test versus red-green discrimination index (RGI) for protan (left panel) and deutan (right panel) subjects. The correlation coefficient for best linear fits is (A) 0.37 and (B) 0.51.

Figs. 2.13 and 2.14 compares the range on the Nagel anomaloscope, defined by RGI (see Equation 2.2), versus the number of errors made on the Ishihara plates (allowing no misreadings) and on the AO-HRR plates. The results show that there is no strong correlation between the Nagel range and the number of errors a subject makes. Fig. 2.13 shows that both protan and deutan subjects can have an RGI value > 0.8 (equivalent to a range < 15 units on the red-green scale) and their performance on the Ishihara test can vary greatly, from making between 2 and 24 errors. Subjects with very low RGI (< 0.2) usually make many errors on the Ishihara plates, in our sample at least 16 errors. And even dichromats can produce between 16 and 24 errors. On the whole, the number of mistakes on the Ishihara test is a very poor predictor of chromatic discriminative ability. Surprisingly, even someone who makes more than 20 errors on the Ishihara can have high discriminative ability, with an RGI value anywhere from 0 to close to 1. These findings illustrate clearly that the diagnosis of dichromatism cannot be made using the Ishihara plates alone.

Fig. 2.14 shows that protan subjects with low RGIs can still make as few as 3 errors and subjects with relatively high RGIs can make up to 9 errors on the AO-

HRR plate test. The correlation between the number of errors on the AO-HRR and the RGI is higher than between the Ishihara and the RGI, but it is still far from being a one-to-one correspondence. Therefore, the number of errors on the AO-HRR also fails to predict adequately the subject's discriminative ability. Again, a diagnosis of dichromatism cannot be made based only on the AO-HRR test.

In summary, for an efficient screening test it is very important to have a high specificity (high PPV) and a relatively high sensitivity (high NPV). This means that there should be no false positive results and it is to a certain extent arguable to accept a moderate number of false negative results (minimal colour deficient observers). The results have demonstrated that the Ishihara and the AO-HRR plates perform best, however neither test can be done alone and additional tests are required to confirm diagnosis. These further tests need to have a high sensitivity and PPV because a positive result will probably lead to diagnosis of colour vision deficiency. These findings suggest that to avoid misclassifying normals as defective, the first screening test should be the AO-HRR, with the highest specificity, and to follow up the negative results with a high sensitivity test such as the Ishihara test (although in practice this would mean performing two tests on most of the population). By using the Ishihara first and following up the positive results with the AO-HRR avoids false negatives.

2.4 Discussion

Current colour vision tests vary in sensitivity and specificity and employ response criteria that make the results difficult to generalise. The findings from this study show that one test alone is not sufficient to correctly assess colour vision. The discrepancies in the results of the various colour vision tests can be partly attributable to the differences in stimulus characteristics and testing criteria for each test. An accurate classification of colour vision therefore involves the use of a battery of tests that can take up to one hour. It is the analysis of the combined results of all the tests that provides the necessary information to assess the correct

diagnosis of colour vision deficiency. The tests compared in this study have very different examination procedures and visual tasks. The tests complement each other but differ significantly in the use of colour signals. The results obtained cannot therefore be expected to show precise agreement.

The Nagel has diagnostic capabilities but the variability of the yellow match parameters is large even within the normal population. The other tests give indications of the severity of the deficiency depending on the number of errors. The results confirm the reliability of the Ishihara plates as a red-green screening test. However it must be used in addition to the AO-HRR to confirm results and if the Nagel anomaloscope is not available the diagnosis of dichromatism cannot be made.

The number of errors made on the Ishihara plates is not a useful guide to the severity of a colour vision defect (CIE 2001b) except that it has been found that only colour normals and a few mild anomalous trichromats make less than about ten mistakes or misreadings on the first 25 plates of the 38-plate Ishihara test (Birch & McKeever 1993; Birch 1997c; Cole 1963).

The lower specificity rate of the Ishihara plates (compared to the AO-HRR, D15 and CU tests) is due to the design of the numerals which may lead to 'misreadings' and previous studies have argued that misreadings should not be included as mistakes or as failure of the plate (Birch 1997c). The results have shown that some normals (on the Nagel anomaloscope) do make mistakes on the Ishihara plates but if misreadings are permitted this would increase the screening efficiency of this test only slightly, to a specificity of 91.7%; see Appendix A for calculation of screening efficiency allowing up to three misreadings and the effect of choosing different pass/fail criteria on the Ishihara as suggested by Birch (1997c).

For example the JAA uses the Ishihara as their first screening test. If applicants make any error on this, they have to go on to a secondary test which they have to pass. Results of a study carried out by Squire et al. (2005) showed that colour vision normals (according to the Nagel anomaloscope) can fail the Ishihara test

and then, in doing a secondary test which involves a lantern test, can fail and not be allowed to become a pilot.

The results for the screening efficiency of the AO-HRR test resembles previous studies (Birch 1997a) that found a screening sensitivity of 98% (401 male colour deficient observers) however their pass/fail limit allowed for one error only on the initial screening plates. In our study, this criterion of allowing one error for passing the test would have lowered the sensitivity by 10% so the AO-HRR criterion of making no errors provides the highest screening efficiency. Other studies (Hardy et al. 1954; Sloan & Habel 1956) have suggested that at least 10-20% of normal trichromats may be expected to fail at least one screening plate of the AO-HRR test. However, the results here show that 5.8% of people with anomalous colour vision can pass this test. This suggests a significant overlap between the responses of normal and colour deficient people which cannot be removed by adjusting the error threshold.

Grading and classification efficiency were not investigated here as various analyses have been performed previously and did not seem relevant to this study. The Nagel anomaloscope is the reference test for grading and classifying and it has been shown that the number of errors on the other tests described in this work do not necessarily relate to severity of deficiency (see Figs. 2.13 and 2.14). For the AO-HRR, it has been shown in a previous study that classification of protan/deutan is uncertain for 18.5% of anomalous trichromats (Birch 1997a), and only two grades of severity can be distinguished with confidence: either slight or severe colour vision deficiency. Failure in classification of the Ishihara plates (Birch 1997c) was shown to be 18.1% for protanopes and 3.2% for deuteranopes who could see neither figure nor the classification plates. Similarly, no classification was obtained for 6.7% of protanomalous and 2.1% of deuteranomalous trichromats. The D15 and CU tests classify correctly only about 20% of colour deficient observers: either type of deficiency is incorrect or uncertain due to an equal number of protan/deutan errors (Birch 1997b).

Another further requirement of a good diagnostic test is that the result is repeatable and is subject to minimal inter-observer variation. This was not

evaluated for the above tests. However, there were subjects attending City University that had previously failed a colour vision test, then passed, and for this reason were referred to us for a more accurate assessment and an explanation of the observed variability in a subject's test results. One subject learnt the number sequence for the Ishihara result and therefore passed the test and others showed variability as a result of chance. These subjects all are minimally colour vision deficient and therefore on some occasions may have enough sensitivity to 'guess' correctly the number shown on the plate.

The other inconvenience about several of these colour vision tests is the lack of diagnostic ability for yellow/blue deficiencies, which are only included in the AO-HRR and the D15. The classification efficiency of these is rather difficult due to the lack of colour deficient observers with this type of deficiency and the inability of the Nagel to assess for yellow/blue. Therefore, the diagnosis of tritan defects is rather problematic and the condition often eludes detection. The tritanopic isochromatic lines converge to a point just outside the blue corner of the chromaticity diagram (see Fig. 1.30), therefore the most appropriate colours for plate tests are blues/greys/yellows. This choice introduces difficulties for the printer because yellows are normally much brighter colours than blues and it is therefore more difficult to achieve with appropriate specification. Special pseudoisochromatic plates have been designed to include plates to assess yellow/blue loss, e.g. the AO-HRR test or by examining chromatic discrimination on pigment arrangement tests, such as the F-M 100 hue test or the D15 (both have a tritan axis). Adaptation tests have also been designed to attempt to detect tritan defects. They are designed to isolate any residual S-cone responses. There is an additional obstacle and this involves distinguishing acquired yellow/blue defects from congenital yellow/blue defects and variations due to inter-observer differences in lens absorption and macular pigment density.

The range of tests described here is not complete, but the tests selected are among the most commonly used. The results obtained bare the current state of colour vision testing and its limitations and emphasises the need for a new test that avoids some of the problems revealed here. The requirement for an accurate colour vision assessment is to have high screening efficiency and ability to grade

and diagnose all types of colour vision deficiencies (including tritan and acquired colour vision).

2.5 Conclusions

Different colour vision tests have been developed by making use of different aspects of visual performance that depend on the use of colour signals. In this chapter several colour vision tests have been described that assess primarily for red/green colour deficiencies. The efficiency of these tests has been compared against the Nagel anomaloscope. The results reflect in some cases the design aim of the colour vision test, i.e. the high screening efficiency of the Ishihara and the AO-HRR pseudoisochromatic plates or the low sensitivity of the D15 for mild deficiencies. None of the tests can provide an unequivocal diagnosis by itself, since there is an overlap between the results of normal and colour vision deficient observers, which cannot be overcome by adjusting the error threshold. Therefore these tests are often used as part of a 'battery' of tests, which combines the results of several tests to obtain a final diagnosis. For an individual subject, the tests can give contradictory results and often a consensus diagnosis may not be easy to reach. The classification of deutan and protan defects is imperfect with the Ishihara and AO-HRR pseudoisochromatic plates and even less precise with the D15 and CU tests. The D15 and the AO-HRR might also identify tritan defects, but this diagnosis is not very reliable. In conclusion, each of the evaluated tests has benefits and specific weaknesses. A more reliable and easy-to-use colour vision test with high screening efficiency and good classification ability for deutan, protan and tritan defects would therefore be useful. The CAD test which is presented and evaluated in the next chapter provides a possible alternative.

3 The CAD test: A New Approach to Colour Assessment and Diagnosis of Colour Deficiency

The methods employed in clinical and occupational colour vision tests, have been described in the previous chapter. These methods tried to ensure that subjects do not see the stimulus by detecting luminance differences between the coloured stimulus and background field. The requirement for accurate assessment of colour vision is to eliminate the use of any luminance contrast signals that may reveal the test figure from the background.

In this chapter, we introduce a new computer based test that avoids most of the problems of conventional colour vision testing. The basic technique employs spatiotemporal luminance noise to isolate the detection of colour signals. In this chapter we describe the full design of the Colour Assessment and Diagnosis test and present data obtained in a large number of subjects to demonstrate its advantages and to produce a template that shows the limits of normal chromatic discrimination sensitivity.

3.1 Introduction

New techniques for colour vision assessment have emerged largely as a result of understanding better the spatiotemporal properties of luminance and chromatic mechanisms and technical advances in high resolution colour displays. A variety of test patterns can be generated and presented under computer control. The subject is normally asked to detect a stimulus that can vary in time and can be controlled in chromaticity along specified axes of colour space. The strength of chromatic saturation can be controlled, thus accommodating for different degrees of deficiency and allowing thresholds to be measured by adaptive methods. Once again the principal requirement is to ensure that observers do not employ luminance or edge cues to detect the stimulus. Several tests have been described previously that make use of colour television to generate chromatic test stimuli (Gunduz et al. 1989; Heard et al. 1987; King-Smith et al. 1989). However, several major problems have confronted the designers of such systems. Reffin et al. (1989) summarised the main disadvantages these designers of computer-display colour vision tests encountered. For example:

- Edge artefacts may reveal the presence of a chromatic boundary either from the slight misalignment of the phosphor guns or from the intrinsic pattern of phosphor dots (Heard et al. 1987).
- If chromaticity differences at high spatial frequencies are presented the chromatic aberration of the eye may render colour boundaries visible.
- Luminance matches vary amongst individuals, both within normal colour vision and colour deficient subjects. In order to ensure isoluminance between the target and the background luminance, flicker photometric matches have to be made prior to the test.

These methods that set the appropriate levels of luminance contrast for each individual observer in a preliminary experiment increase the complexity of the experiment and prolong the testing time. Similar problems faced the designers of the first pseudoisochromatic plates. Characteristic colour confusions found in protan, deutan and tritan defects have been exploited in clinical colour vision tests and failure to equate effectively luminance contrast differences may lead to errors

in the design of pseudoisochromatic plates. The main requirement is to produce a recognisable pattern using colours within a given isochromatic zone without introducing significant luminance contrast signals to enable detection of the test pattern by the luminance system of a particular class of observer. The relative luminous efficiency of the eye varies considerably within normal trichromats, and even more so in protanopes and deuteranopes. Isochromatism therefore exhibits considerable variation amongst normal and colour deficient observers. Even if one particular design was suitable for all subjects, it would be difficult to implement it in practice since accuracy depends on the quality of the printing and on the illumination used (Lakowski 1965b; Lakowski 1966), i.e. minor misalignments between the figure of one chromaticity and a uniform background of another undistinguishable chromaticity, created as a result of the printing process, may reveal the shape of the figures. The designers of the AO-HRR and the Ishihara pseudoisochromatic plates have minimised these problems by breaking up the test figure and the background into a number of discrete patches, each with its own shape and contour. Additionally, the luminance of the patches varies randomly rather than being equated. The AO-HRR test plates use coloured targets embedded in an achromatic background (see Fig. 2.2), while the Ishihara plates use a chromatic difference within the figure in comparison to the background causing some subjects to see different figures (see Fig. 2.1).

A measure of chromatic sensitivity is difficult to extract from standard colour vision tests. Most colour vision tests represent working compromises that limit their usefulness for screening and detecting and do not therefore give precise indices for colour sensitivity loss. In some cases the outcome is dependent on motivational, attention and learning factors. Not only is it a necessity to have a colour vision assessment that has a high screening efficiency and is reproducible with small variance but that can also yield a quantitative assessment of the severity and type of the deficiency.

3.2 Methods

All visual stimuli used in our studies were generated on stable cathode-ray tube (CRT) displays. A CRT monitor provides the means to produce spatiotemporal modulations of light intensity and spectral content and this is useful when one needs to stimulate selectively the cone photoreceptors in the eye. The Colour Assessment and Diagnosis (CAD) test is based on a spatiotemporal masking technique that isolates the use of colour signals and a visual psychophysical procedure for measuring thresholds for a colour response.

3.2.1 Spatio-temporal luminance noise

The CAD test was designed to overcome some of the difficulties associated with conventional colour vision tests by employing a spatiotemporal luminance contrast perturbation technique. A design, similar to the original pseudoisochromatic plates is buried into an array of achromatic checks. The luminance of each check varies randomly with equal probability within a specified percentage of background luminance. Although there are local increments and decrements in luminance, the luminance of the checks are such that at any point in time the space-averaged luminance over the whole array is constant and equal to the background luminance. This random luminance modulation (RLM) goes on independently of any chromatic displacement involved. Fig. 3.1 shows the luminance values for horizontal cross sections through the stimulus array. The RLM of each frame can be specified both spatially and temporally by the examiner. The amplitude of spatial RLM determines the level of luminance masking achieved.

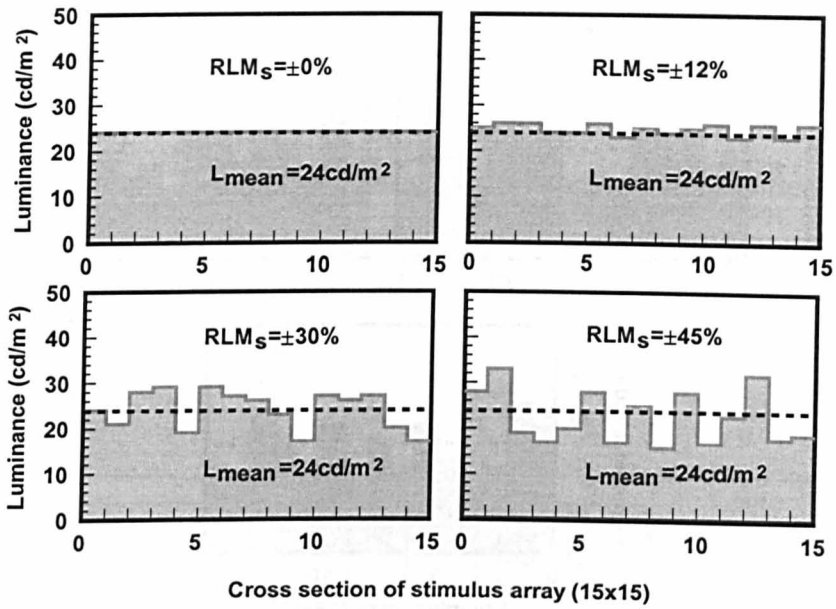


Figure 3.1: Diagram showing luminance profiles for individual frames of spatial random luminance modulation (RLMs) for the stimulus array (15 by 15 checks). Profiles are shown for screen background level of 24 cd/m^2 for different RLMs amplitudes. The mean background luminance is shown by the black dotted line.

Temporal RLM can be used alone or together with spatial RLM. When specified in conjunction the effect is to create a condition called *dynamic luminance contrast (LC) noise*, otherwise the condition is *static LC noise.* Dynamic RLM, causes the perception of local flicker and random motion of the checks. For dynamic LC noise the luminance of each check varies at a frequency of 20 to 12.5 Hz. Fig. 3.2 shows an example of dynamic LC noise used in this study for evaluating chromatic sensitivity.

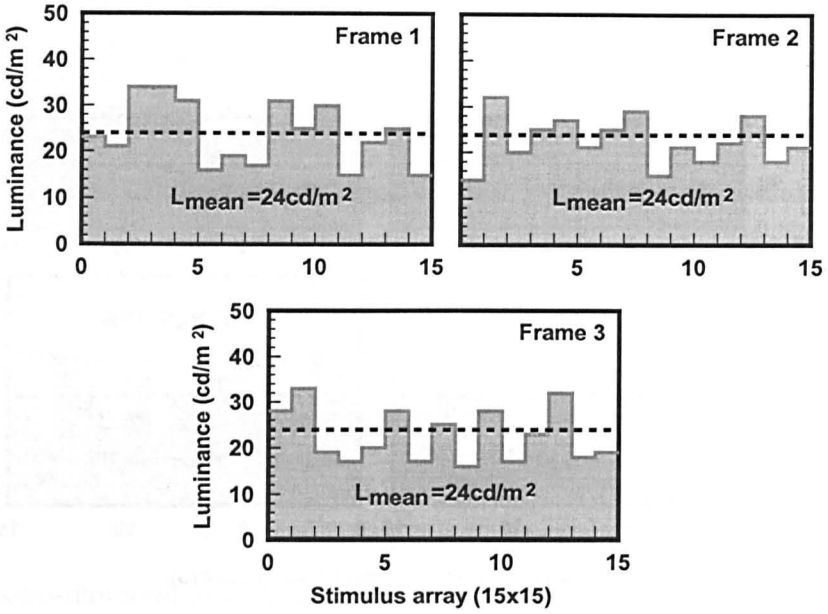


Figure 3.2: An example of dynamic luminance contrast noise. In this example the mean background luminance remains constant at 24 cd/m^2 throughout the stimulus array and in time, while the individual checks vary in luminance with each frame within a specified level of ±45% of the mean background luminance.

The principle of this method is based largely on the spatial (Barbur and Ruddock 1980) and the temporal (Barbur et al. 1981) background perturbation techniques that have been developed to investigate the spatiotemporal properties of human vision. Previous studies have made use of either one of these principles (Regan et al. 1994) or both (Barbur et al. 1991; Barbur et al. 1993; Barbur et al. 1994; Barbur et al. 1997; Birch et al. 1992) to generate psychophysical tests for colour discrimination.

3.2.2 Effectiveness of the luminance contrast technique

When chromatic signals are introduced during the time course of LC noise, the subject can only make use of chromatic signals, provided the LC noise is sufficiently high. To verify the efficiency of this technique, the effect of RLM amplitude on detection of an achromatic stimulus pattern was investigated and is

shown in Fig. 3.3. The effect of dynamic compared to static LC noise on masking the detection of achromatic motion thresholds is much more efficient when dynamic LC noise is present. The two examples shown indicate subjects at the extreme of the range. The difference in gradient has not been investigated further and may well be related to the relative combination of magnocellular and parvocellular pathways to the detection of the dynamic LC noise.

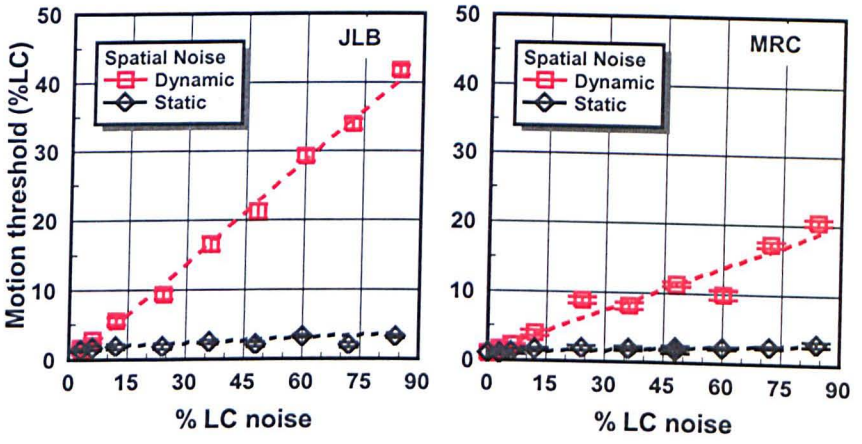


Figure 3.3: The effect of RLM amplitude on thresholds for detection of an achromatic stimulus pattern (similar to Fig. 3.7) in the fovea. When the LC noise varies dynamically the detection of the stimulus rises linearly with the amount of LC noise.

On the other hand, Fig. 3.4A shows that the dynamic and static noise does not affect significantly the subject's ability to discriminate a chromatic difference. However, if one examines a colour deficient observer the results show how chromatic thresholds increase for increasing levels of LC noise revealing the particular colour deficiency of the observer (see Fig. 3.4B).

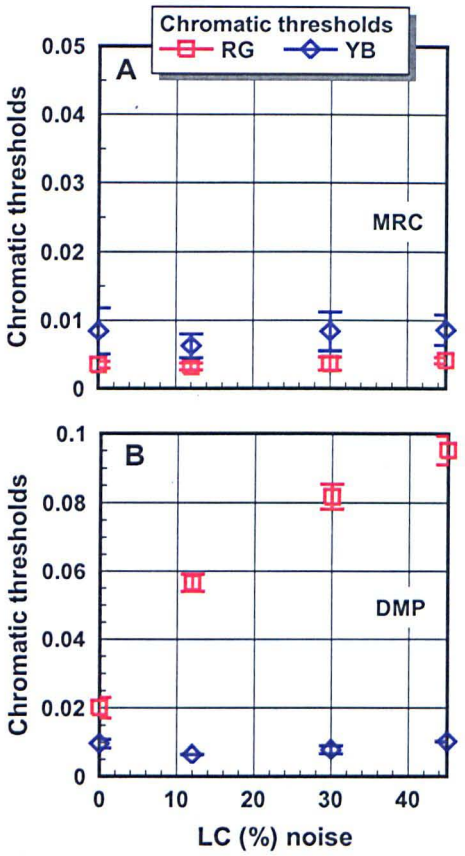


Figure 3.4: The effect of dynamic LC noise amplitude on thresholds for the detection of colour-defined moving stimuli (shown in Fig. 3.7). Chromatic detection thresholds for the red-green (RG) and yellow-blue (YB) channels are plotted. Thresholds are computed by averaging eight directions in colour space for each colour channel (see Section 3.5 for a description). (A) Data for a normal trichromat. Chromatic detection thresholds are invariant for increased LC noise amplitudes. (B) Data for a deuteranopic observer. RG detection thresholds increase as the level of noise increases, whereas YB thresholds remain invariant for different percentages of LC noise levels.

Fig. 3.5 and 3.4B show results for the same observer plotted on the 1931 CIE chromaticity chart. As the LC noise is increased, the observer is no longer able to make use of residual luminance contrast signals and his thresholds increase monotonically with LC noise. The largest chromatic displacements away from background chromaticity are set by the isoluminant condition and the limits imposed by the phosphors of the display.

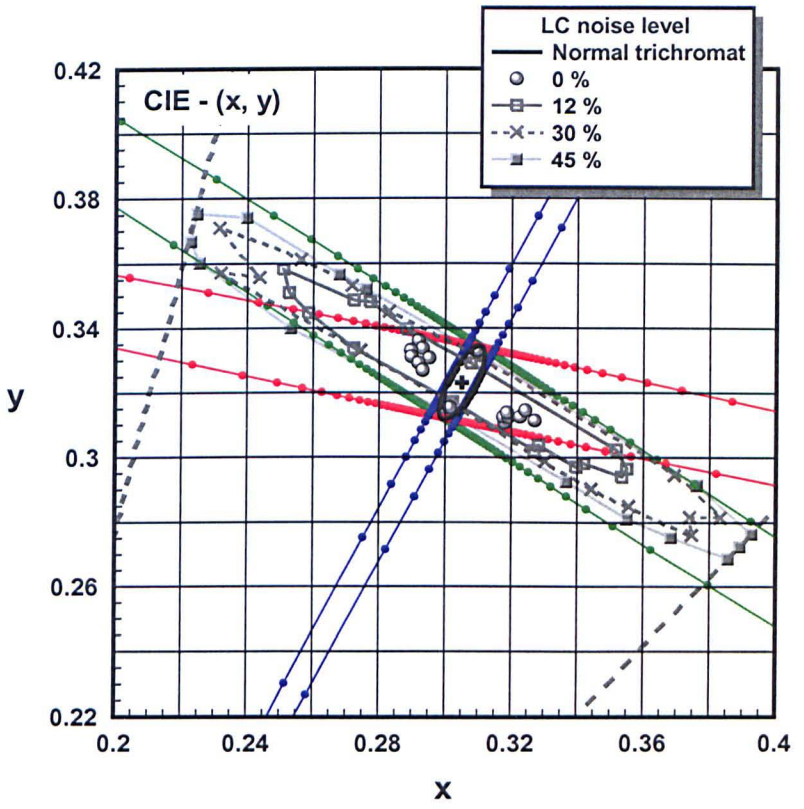


Figure 3.5: Chromatic threshold contours for a deuteranopic observer are shown on the CIE-x,y chromaticity chart. The ellipse for a normal trichromat is shown as a solid black contour. In the absence of LC noise (circles); the subject's thresholds for YB discrimination may be smaller than those measured in a typical normal trichromat. The deuteranope can make use of residual luminance contrast signals to detect the test stimulus, more so than a normal trichromat since the test stimulus is "isoluminant" for the standard CIE normal observer. His RG thresholds (particularly along a deutan isochromatic line) are increased. The broken lines show the phosphor limits for 45% modulation. For RLM amplitudes >45%, this subject reaches the limit of chromatic displacement set by the phosphors of the display without the subject being able to detect any colour signals.

The results reported here demonstrate how the method can be used for measurement of discrimination ellipses and thresholds in normal trichromats, similar to those reported by MacAdam (1942) and isochromatic zones in dichromatic subjects. This technique has been used in other studies to show how the pupil responds selectively to the colour stimuli embedded in LC noise. The pupillary responses are totally absent when the colours selected fall within the corresponding dichromats isochromatic zones (Barbur 2004). An objective colour vision test based on pupillary responses to chromatic stimuli along characteristic isochromatic lines found in protanopia, deuteranopia and tritanopia has been previously described (Barbur et al. 1993; Saini and Cohen 1979; Young and

Alpern 1980). The brief presentation of a spatially uniform, coloured stimulus or an isochromatic grating produces a transient constriction of the pupil which varies systematically in amplitude, latency with chromatic contrast (Rodriguez-Carmona et al. 2003) and grating spatial frequency (Barbur 1991).

The spatiotemporal perturbation studies described here show how this technique can be applied to the measurement of red-green (RG) and yellow-blue (YB) chromatic sensitivity in human vision.

3.2.3 Equipment

The stimuli are generated on a high resolution, 21-inch Sony Trinitron monitor (model Multiscan 500PS, 1280 x 1024 pixels, operating at a frame rate of 60 Hz, 20 cd/m^2 maximum luminance), driven by an ELSA Gloria XL 10 bit graphics card. Control programs ran on a PC under MS-DOS rather than Windows to ensure reliable accurate timing and real time operation. Observer responses were recorded via a button box attached to an input/output board (Amplicon PC30AT). Programming was carried out by Alister Harlow using C++.

Figure 3.6: Photograph of experiment set-up. The headrest facilitates fixation and ensures the viewing distance is 700 mm. The response button box is also shown; the subject presses one of the four buttons according to the direction of motion of the stimulus.

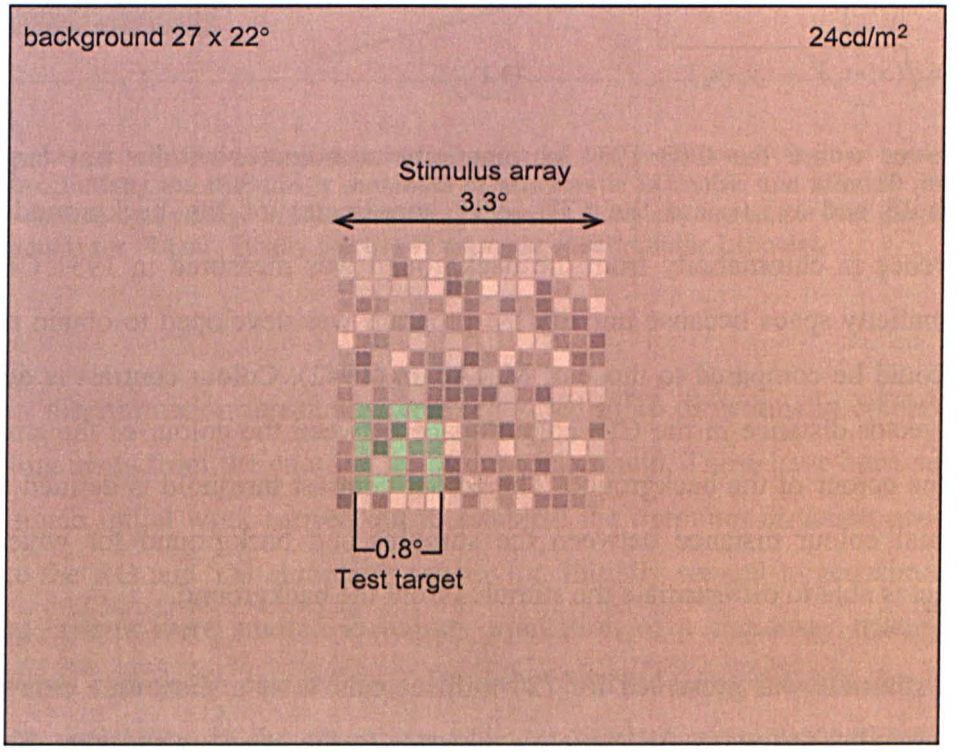


Figure 3.7: Schematic of the stimulus used to measure chromatic thresholds showing a square array of 15 x 15 checks, subtending 3.3°. The test target subtends 0.8° and is defined by having 21 coloured checks in a 5 x 5 square array. The background luminance was set at 24 cd/m^2.

The stimulus was generated at the centre of a uniform background field subtending an angle of 27° by 22°. The stimulus consists of checks arranged in a matrix array and subtends a visual angle of 3.3° by 3.3° at a viewing distance of 700 mm. The uniform background field is set at a luminance of 24 cd/m^2. Fig. 3.7 shows an example of this stimulus. The luminance of each check changed randomly every 53 to 80 milliseconds and took values within a range of ±45% of background luminance. The background chromaticity was chosen to correspond to the 'white' reference used by MacAdam (1942) in 1931 CIE-x,y chromaticity coordinates; x=0.305, y=0.323. The checks forming the coloured test target are arranged as a square 5 by 5 checks and consist of 21 colour-filled checks. The change in chromaticity of the test target is in any one of the number of specified directions away from the background chromaticity towards the spectrum locus. Chromatic difference (CD) is measured as the distance in the 1931 CIE-x,y chromaticity chart away from the background chromaticity and is controlled by the randomly-interleaved, multiple-staircases (one staircase for each direction of colour space tested).

$$CD = \sqrt{(x_t - x_b)^2 + (y_t - y_b)^2} \qquad (3.1)$$

where x_t, y_t are the CIE 1931 chromaticity coordinates of the test target at threshold and x_b, y_b are the CIE 1931 coordinates of the background. The difference in chromaticity from the background was measured in 1931 CIE-x,y chromaticity space because initially the program was developed to obtain results that could be compared to those of MacAdam (1942). Colour contrast is defined as a vector distance in the CIE colour space between the colour of the stimulus and the colour of the background. The colour contrast threshold is defined as the minimal colour distance between the stimulus and background for which the subject is able to differentiate the stimulus from the background.

Each stimulus was presented for 720 milliseconds, after a short time delay (160 milliseconds) following the subject's response to the previous stimulus. Fig. 3.8 shows an example of the time sequence of stimulus presentation. The RLM was present throughout the stimulus presentation time. The colour signal strength needed for threshold detection of the moving test stimulus for each hue direction

investigated is measured using a four alternative-forced-choice (AFC) procedure, that yields a low chance probability of 1/16 (Barbur 2004). The subject's task is to indicate the direction of movement of the colour-defined stimulus by pressing one of four corresponding buttons on the response box. If no test target is seen the subject is asked to 'guess' and press any of the four buttons.

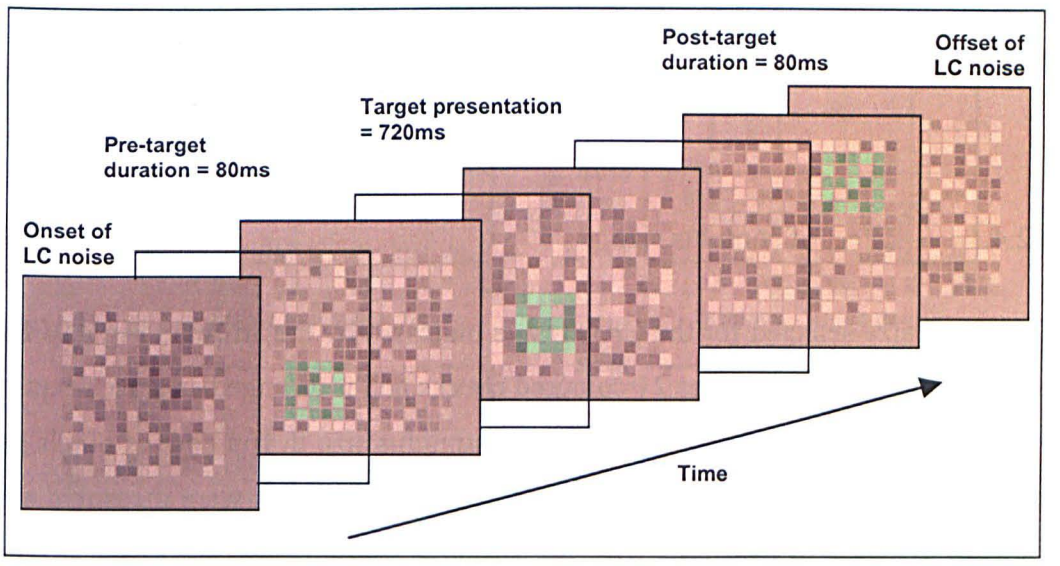


Figure 3.8: A schematic representation of some of the sequence of frames used in the stimulus. Initially the stimulus is composed of achromatic LC noise, and after 80 ms a test target defined by a change in chromaticity that moves along any one of the four diagonals of the stimulus for 720 ms. Finally there is 80 ms again of achromatic LC noise.

Colour discrimination thresholds were measured in 16 directions in colour space radiating away from the coordinates of the background. These have been selected after much initial work carried out to establish the optimum direction needed to isolate the RG and YB chromatic sensitivity. Initially several target stimuli and testing criteria were studied including orientation of a stationary triangle. No significant changes in colour thresholds were observed using the different stimuli but the variability of the measurements was smaller when using the moving stimulus configuration.

The observers head movements were restricted by use of a chin rest and forehead support as shown in Fig. 3.6.

3.2.4 Definition of stimulus parameters

In order to ensure a 'stable' measure of chromatic sensitivity, the effect of stimulus size and background luminance level was investigated for a small number of observers.

3.2.4.1 Effect of stimulus size

The target size selected for the CAD test is indicated by the arrow in Fig 3.9. The optimised target stimulus has 21 coloured checks and is arranged as a 5 x 5 square matrix. In order to investigate the effect of stimulus size the number of pixels per check was altered from 4 to 14. The RG and YB thresholds were measured at a viewing distance of 700 mm with the stimulus presented at the fovea. As the stimulus size is increased, the thresholds for chromatic sensitivity reach a plateau and the thresholds no longer change significantly. The choice of 0.8 square degrees for the target size ensures no significant improvements in thresholds result by increasing the stimulus further (indicated by the grey arrow in Fig. 3.9).

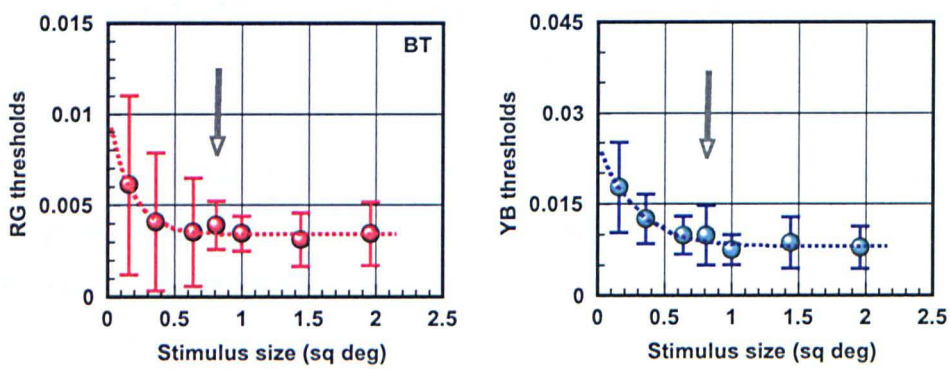


Figure 3.9: Continues on next page.

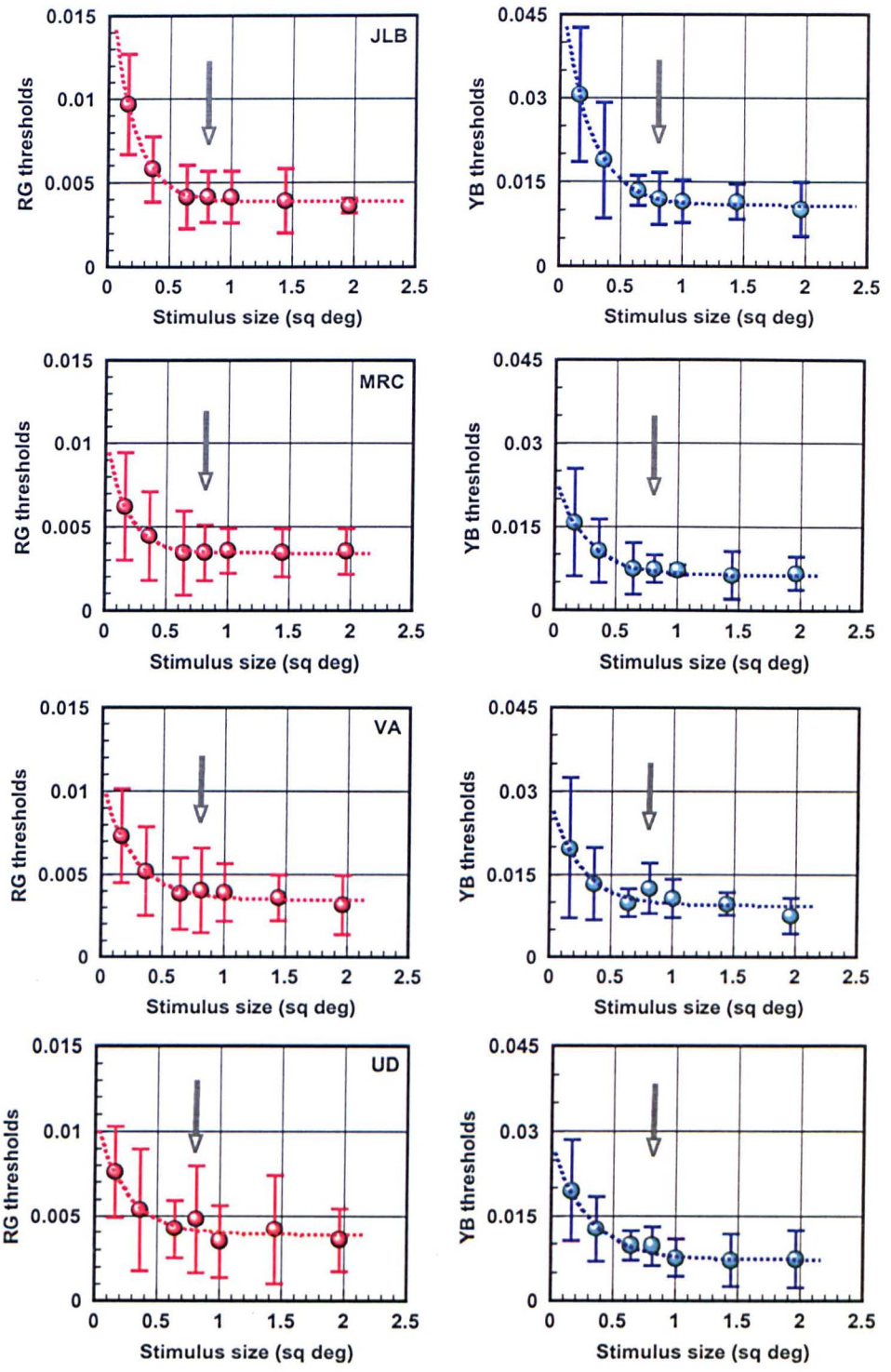


Figure 3.9: Foveal RG (left) and YB (right) chromatic thresholds measured for different target sizes for 5 observers. The target stimulus size varied from 0.16 to 1.96 sq degrees (4 to 14 pixels per check). The arrow indicates the size of stimulus chosen for the measurements carried out in this study. The error bars show +/- 2SD. The best fit line ($x_1 + e^{(x_2 \cdot size)} + x_3$) has been included for representation purposes.

3.2.4.2 Effect of background luminance

The effect of the background luminance was varied from 52 to 0.05 cd/m^2 in order to investigate the effect of background luminance on chromatic thresholds. Fig. 3.10 shows the results obtained and the arrow indicates the background luminance selected for the CAD test.

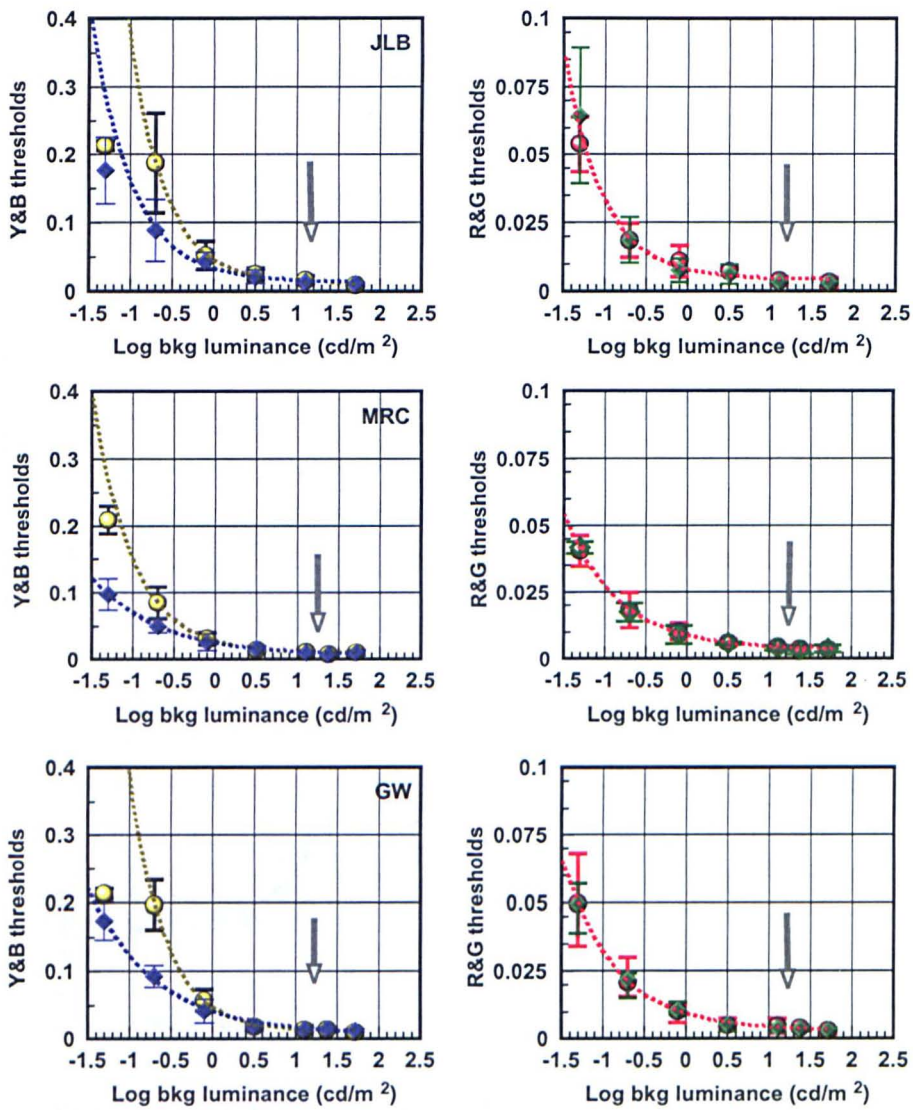


Figure 3.10: Continues on next page.

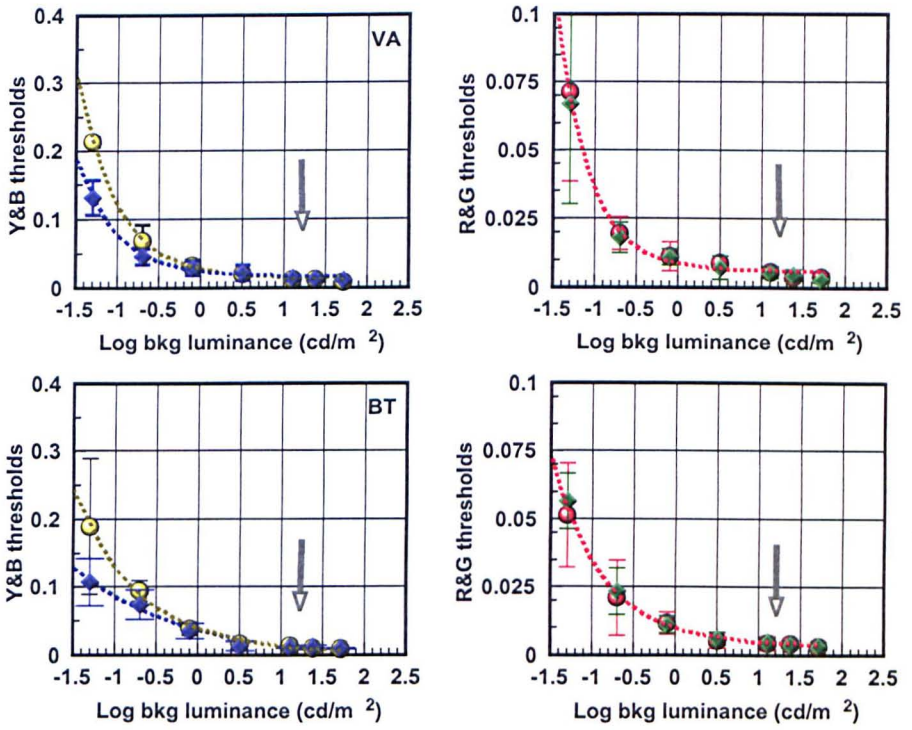


Figure 3.10: Foveal R, G, Y and B chromatic thresholds measured for different background (bkg) luminance levels plotted on a log scale for 5 observers. The background luminance was varied from 0.05 to 52 cd/m^2 (-1.3 to 1.72 on a log luminance scale). Neutral density filters were used for the lowest three values of luminance ensuring the correct reproduction of chromaticities whilst allowing to reduce the background luminance to low levels. The background luminance chosen for the measurements carried out in this study is 24 cd/m^2 which is indicated by the arrow. The error bars show +/- 2SD. Dotted lines represent best fit lines ($x_1 + e^{(x_2 \cdot bkg)} + x_3$) drawn for graphical purposes.

The results show the background luminance level has a significant effect on chromatic sensitivity. The increase in colour thresholds was particularly evident below 1 cd/m^2 (0 on a log scale, as shown in Fig 3.10). For this reason, chromatic thresholds are shown separately for the yellow (Y), blue (B), red (R) and green (G) directions in colour space. This is consistent with previous studies (Brown 1951) that have found a non-uniform increase of chromatic thresholds with background luminance at the fovea. These findings are also consistent with the rise in tritan errors for the F-M 100 hue test when performed under low light levels (Bowman and Cole 1980; Knight et al. 1998; Knoblauch 1987; Smith et al. 1991). The findings from this study have shown that there is a greater reduction in sensitivity in the Y compared to the B direction observed at the lower values of

background luminance. The observed asymmetry in chromatic sensitivity could be due to the unequal excitations produced in S-cones when presented with increment or decrement S-cone signals at low light levels (Yeh et al. 1993). The advantage of choosing 24 cd/m^2 as the background luminance for the CAD test not only ensures symmetry in chromatic thresholds but also avoids any changes in chromatic sensitivity as a result of signal changes in retinal illuminance that may arise due to fluctuations in pupil size.

Similar data could be produced for any eccentric location. Since the purpose of the CAD test is to diagnose and quantify colour vision it is intended that the CAD test is carried out in the fovea.

3.2.5 Staircase method

A multiple staircase method is used in these experiments to determine accurately chromatic discrimination thresholds. The staircase procedure is often used in experiments in which visual threshold stimuli are to be determined accurately in a relatively short time. It is an example of adaptive testing procedures used to keep the test stimuli close to the threshold by adapting the sequence of stimulus presentations according to the observer's response. In each staircase, testing begins at a certain hue saturation and proceeds to a less saturated hue every time the subject makes a 'correct' response twice in a row. One 'incorrect' response is followed by presentation of hues with higher saturation value. The staircase parameters are given in Table 3.1 and have been chosen from preliminary experimental runs. The staircases are randomly interleaved thus reducing the effects of series-interdependencies biasing that may arise from learning the procedure (Cornsweet 1962). The step-size is variable during the course of the experiment; initially the step is largest (0.025) and reduces in size exponentially to a final step-size of 0.001, in order to maximise efficiency of the measurement procedure. Testing on any one staircase was terminated after 12 reversals and the mean of the last six reversals was taken as the threshold estimate for the colour

direction tested. The value obtained from each staircase gives the chromatic threshold measured for one colour direction.

Staircase Parameters	
Initial value	0.025
No. of reversals	12
Reversals ignored	6
Start increment	0.025
End increment	0.001

Table 3.1: Staircase parameters used to measure threshold discrimination parameters.

3.2.6 Procedure of testing

The subject is asked to view the display binocularly from a distance 700 mm. The position of the head and the height of seat was carefully adjusted. In order to explain the testing procedure, the subject was presented with a simplified version of the CAD test lasting only 90 seconds. During that time the test target moves without interruption randomly along the four diagonals of the stimulus. The chromaticity of the target changes in time by following a concentric ring set at 0.078 CD units from the background chromaticity, as shown in Fig. 3.11. This is called the rapid CAD 'dichromatism test' as it identifies only moderate and severe colour deficient subjects. The uncalibrated web version of this test was used in an extensive investigation by Seshadri (2005).

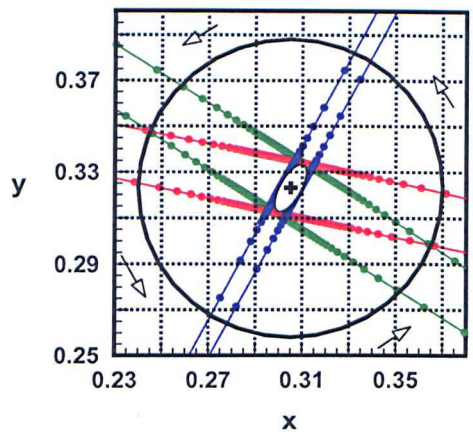


Figure 3.11: Schematic of the rapid dichromatism test. The chromaticity of the coloured target follows the path shown by the black circle specified at a chromatic displacement of 0.078 from background chromaticity points (0.305, 0.323). The dichromatic confusion lines are shown in red, green and blue for the protanope, deuteranope and tritanope, respectively (see Fig. 1.29).

The outcome of the rapid dichromatism test informs of the severity of the colour deficiency. If the subject fails to see the coloured target at any point as it rotates in chromaticity during the presentation this suggests either moderate or severe colour vision loss. The location in colour space where the target was not seen gives an indication about the type of deficiency. This then allows a more specific recipe to be loaded for the complete version of the colour vision test with the colour directions concentrated close to the particular dichromatic colour confusion band. If the subject sees the coloured target all the time during the presentation this suggests either mild colour deficiency or normal colour vision. We have included events during the test when the coloured target disappears for all observers to ensure the subject understands the instructions of the test and does not cheat. Following this rapid test one carries out a more accurate colour assessment that involves specific 'recipes' for normal, protan and deutan subjects. These recipes test colour vision thresholds along 16 directions in the CIE-x,y colour chart especially chosen along the RG and YB axes of confusion for both colour deficient and colour normal subjects. The 16 directions selected allow isolation of red/green and yellow/blue deficiencies and make it possible to distinguish between minimum protanomalous and deuteranomalous deficiencies. For more severe colour deficient observers, those that fail the dichromatism test, chromatic

displacement was measured along the most affected colour confusion axis, i.e. either along the deuteranopic axis for deutan defects or the protanopic axis for protan defects. We have not tested many tritan colour deficient observers due to the low prevalence of this type of colour vision deficiency. The directions in colour space measured were adopted on the basis of earlier measurements and their validity was tested empirically in the course of the present experiments. For the *deutan* and *protan* recipes shown in Table 3.2, six directions with the highest CD were averaged in total along the particular confusion axis (three towards the 'reddish' and three towards the 'greenish' extremes of the colour space diagram) and two directions for the YB direction. For the *normal* recipe (Table 3.2), eight directions in colour space are averaged for each of the RG and YB directions. For the YB direction 56, 60, 64, 68, 236, 240, 244, 248 are averaged and for the RG direction 140, 145, 170, 175, 320, 325, 350, 355 are averaged. More accurate thresholds were computed this way compared to using an ellipse-specific algorithm. Initially, an algorithm such as the one by Fitzgibbon et al. (1999) was used. However, it was found that sometimes miscalculation of the major axis would occur depending on the spatial distribution of the data points.

Type of recipe (angles in colour space)					
Normal		Deutan		Protan	
56	236	65	245	2	186
60	240	135	316	78	258
64	244	138	320	158	342
68	248	142	323	165	346
140	320	147	326	167	348
145	325	152	329	172	352
170	350	155	332	177	356
175	355	162	336	179	358

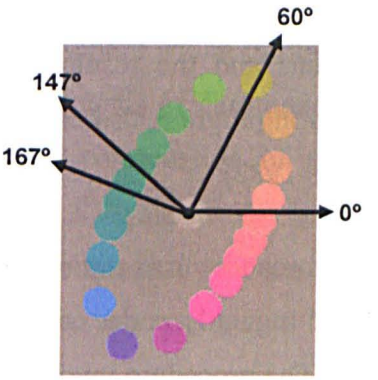


Table 3.2: Angles in degrees used for testing normal trichromats, deutan and protan observers; angles measured in anticlockwise starting from the abscissa.

The spatial arrangement of the four response buttons corresponds to the four possible directions of motion of the coloured target along the diagonals of the square stimulus. The subject's task was simply to press the appropriate button following the presentation of the target to indicate the correct direction of movement. During the test, trials on each direction were randomly interleaved and the chromatic difference between target and field along a given colour direction was adaptively increased or decreased according to the subject's responses on the previous presentation of the same direction. A small number of presentations (5%), intermingled with test presentations, served as control trials. On these trials a target was presented at well above threshold and these served to reinforce the task. The average time for the test is ~ 12 minutes (with no intermediate break).

3.2.7 CRT monitor calibration

In order to generate stimuli with specific chromaticity and photopic luminance on a CRT monitor, it is necessary to know accurately the chromaticity of the red, green and blue phosphors, and the relationship between electron gun voltage and luminance for the red, green and blue guns. The software (written by Alister Harlow) includes a program for automatic calibration of phosphor luminances, so that the relation between the bit values set on the driver card for the red, green and blue guns and the resulting screen luminance are precisely determined. The calibration program also contains values for the CIE standard observer (Wyszecki & Stiles 1982).

The luminance characteristics of the monitor were calibrated automatically and updated regularly under computer control using a luminance meter (LMT, model 1000). The procedure involved determining the luminance versus applied voltage relationship for each gun. Measurements were acquired every 2 voltage steps of the 1024 steps available for the 10-bit graphics card (2^{10}) (Fig. 3.12B). The spectral output of each phosphor was measured using a Gamma Scientific telespectroradiometer (Model 2030-31) and provided the chromaticity co-ordinates of each phosphor. Radiance data was acquired at 5 nm wavelength

intervals from 380-780 nm, with each gun set at maximum voltage. Fig. 3.12A shows the spectral radiance distribution for the monitor used and Fig. 3.12B shows the luminance versus gun voltage for the three guns. The chromaticities of the three phosphors are plotted in Fig. 3.13 and were computed from the phosphor radiance data according to the following equations:

$$X = k\int L_{e\lambda}\bar{x}(\lambda)d\lambda$$
$$Y = k\int L_{e\lambda}\bar{y}(\lambda)d\lambda \qquad (3.2)$$
$$Z = k\int L_{e\lambda}\bar{z}(\lambda)d\lambda$$

$$x = \frac{X}{(X+Y+Z)}$$
$$y = \frac{Y}{(X+Y+Z)} \qquad (3.3)$$

where x,y are CIE 1931 chromaticity coordinates, X, Y, Z are the CIE 1931 tristimulus values, k a normalising factor, $L_{e\lambda}$ phosphor radiance, and $\bar{x}(\lambda),\bar{y}(\lambda),\bar{z}(\lambda)$ are the CIE 1931 colour matching functions. For the three primary display system of the CRT monitor, stimulus chromaticity is restricted to a gamut of chromaticities, which at its largest consists of a triangular shaped region bounded by the chromaticities of the three phosphors which diminishes in size with increasing stimulus luminance. Within the bounds of chromaticity and luminance set by the monitor's characteristics, any combination of photopic luminance and chromaticity can be generated. Fig. 3.13 shows the gamut of chromaticities reproducible on the monitor employed in this study for a maximum luminance of 24 cd/m^2 ±45%.

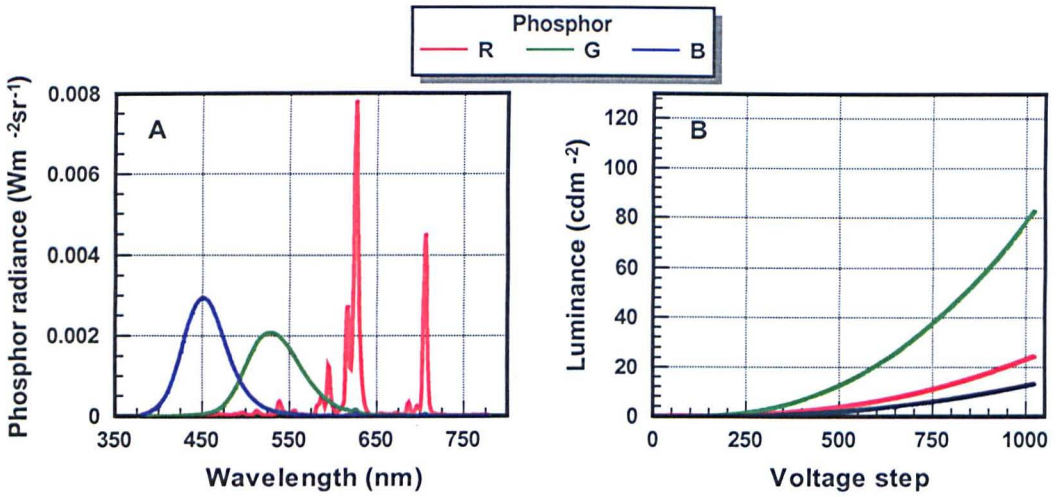


Figure 3.12: (A) Relative intensity output for the red (R), green (G) and blue (B) phosphors of the CRT display. (B) Luminance versus gun voltage calibration for the R, G and B phosphors.

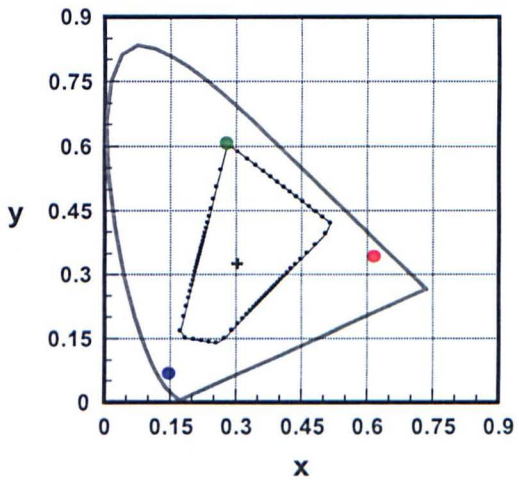


Figure 3.13: Example of the region x,y-chromaticity space that can be reproduced on a typical CRT monitor (shown by the dashed line) plotted in the 1931 CIE-x,y chromaticity diagram. The phosphor limits are calculated for 45% luminance contrast noise. The chromaticities of the three phosphors are also plotted.

The luminance versus gun voltage relationship of a CRT may change within time depending on how often the monitor is used. The luminance versus gun voltage relationship was therefore calibrated frequently during the study. The system was turned on at least 45 minutes before testing to allow for warming-up.

3.2.8 Subjects

Chromatic sensitivity has been measured with the CAD test in a total of 472 subjects. Table 2.2 shows the number of subjects tested on each colour vision test, such as the Nagel anomaloscope, the Ishihara plates, the AO-HRR, the D15 and the CU tests. For more detail see Section 2.3.1.1.

3.3 Results

3.3.1 Typical congenital patterns of colour deficiency

The validity of the measures of chromatic sensitivity using the technique that we have described depends critically on the spatiotemporal LC noise raising the threshold for the detection of LC selectively and leaving chromatic sensitivity unaffected. Fig. 3.14 reveals that the choice of directions tested is appropriate for testing typical congenital colour vision deficiencies. The directions for the deuteranopic and protanopic axes differ (see Fig. 3.5). It is worth mentioning that the difference lies in the orientation of the YB chromatic channel between deuteranopes and protanopes as a result of the underlying visual pigments present. Anomalous trichromats could theoretically have any intermediate orientation. Data shown in Fig. 3.14 suggest that this is however not the case and although a range of orientations is found, the majority of anomalous observers have highest thresholds approximately around the centre of the range of angles tested for the dichromatic axes. No accurate conclusions can be established as the variability observed in the YB chromatic channel amongst normal trichromats is of the order of ± 7° (see Section 4.3.1). It is difficult, therefore, to determine with certainty the orientation of maximum chromatic threshold. Nonetheless, the angles chosen distinguish effectively between protan and deutan colour deficient observers.

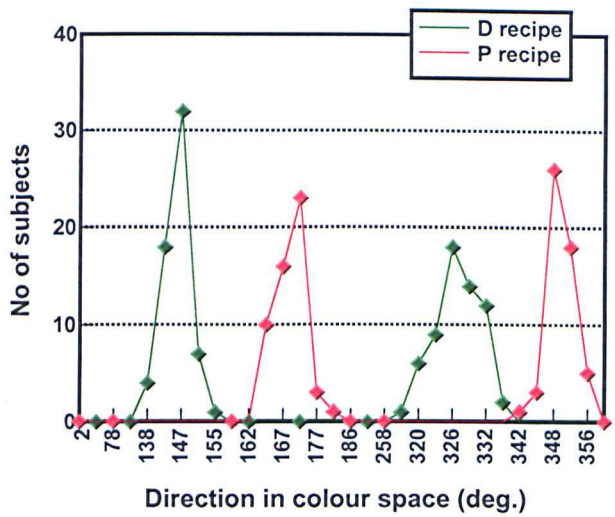


Figure 3.14: Frequency plot showing the number of subjects against the direction in colour space (in degrees) with the highest RG chromatic thresholds. For those subjects that carried out the deutan (D) recipe, the most frequent direction having the highest thresholds was 147° and 326°. The majority of subjects that carried out the protan (P) recipe had peak thresholds aligned at 172° and 348°.

Figs. 3.15 and 3.16 are an example of the results obtained in several typical congenital colour vision deficient observers. The type and extent of deficiency is indicated by the spread along a particular colour confusion band.

Fig. 3.15 shows results obtained from protan subjects. Three of the subjects failed the rapid 'dichromatism' test. Subject D performed the *normal* recipe (see Table 3.2) as a result of passing the rapid 'dichromatism' test and his threshold for the protanopic direction is 28 ($x10^3$) units in CIE-x,y space. For convenience of representation the colour thresholds are represented as the actual distance in the CIE-x,y diagram multiplied by 10^3. This subject has a Nagel RGI of 0.82 with a midpoint of 58, and makes 14 errors on the Ishihara plates, only one error on the AO-HRR and passes both the D15 and CU tests. For subject A the RG threshold on the CAD test is of the order of 88 units in the CIE-x,y space, the Nagel RGI is 0 (whole range) and fails the Ishihara, AO-HRR, D15 and CU tests.

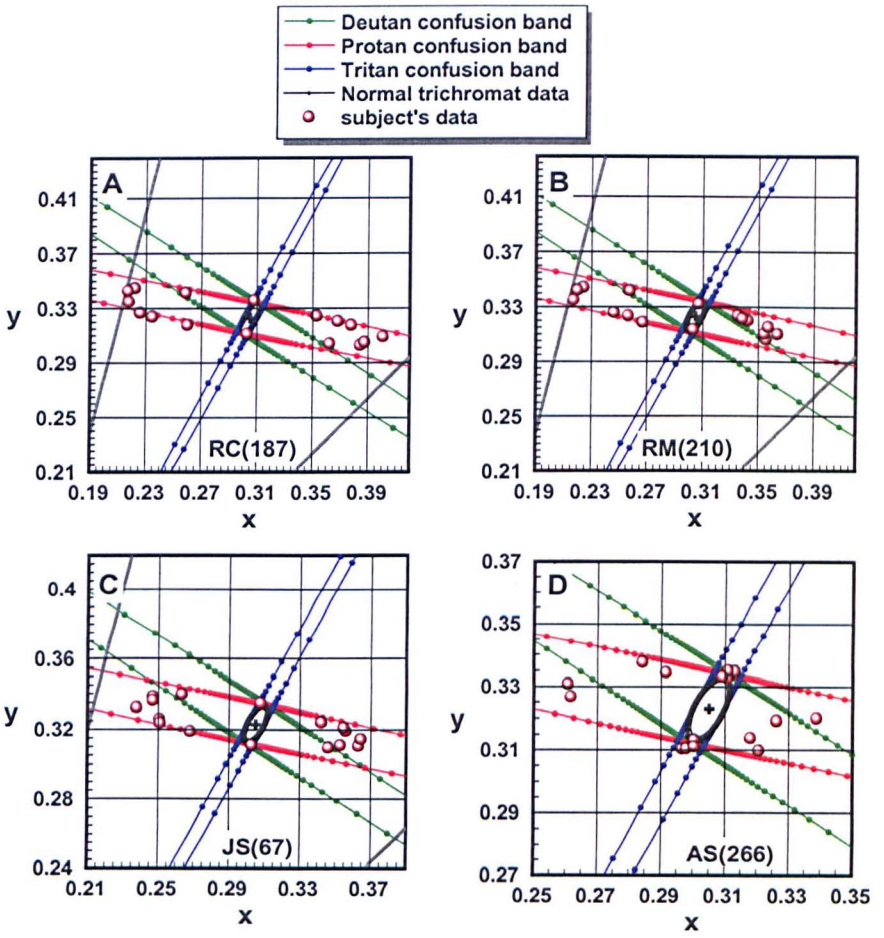


Figure 3.15: Chromatic threshold contours for four protan observers ranging in the severity of their colour vision loss are plotted on the CIE-x,y chromaticity chart. The data for a normal trichromat is shown as a black contour. The largest chromatic displacements away from background chromaticity, as set by the isoluminant condition and the limits imposed by the phosphors of the display, are shown as grey lines. The extent of colour vision loss is related to the elongation along the protanopic confusion band and suggests that the greater the elongation, the lower the level of chromatic sensitivity.

			CAD test		Colour Vision test					
Subject	M/F	Age	RG threshold	YB threshold	Nagel range	RGI	Ishihara	AO-HRR	D15†	CU
A	M	24	88	12	0-73	0	22	10	F	8
B	M	55	71	10	0-63	0.14	21	8	F	6
C	M	48	60	12	35-56	0.70	23	8	F	3
D	M	23	28	13	52-64	0.82	14	1	P	0

† F = fail, P = pass

Table 3.3: Summary of results for subjects A-D in Fig. 3.15. For all four subjects the RG and YB thresholds on the CAD test have been computed. The results on the Nagel anomaloscope: red-green range and RGI, the Ishihara (number of strict errors, i.e. no allowance for misreadings), the AO-HRR and the CU tests are given for each subject. For the D15 test a pass (P) or fail (F) is indicated (same P/F criteria as in Section 2.3.1).

Fig. 3.16 shows typical results obtained from deutan subjects. Subjects E and F failed the 'dichromatism' test and went on to do the deutan recipe. Subject E shows chromatic thresholds along the deuteranopic axis that extend to the phosphor limits. Subject E's Nagel colour match is from 0-73 on the red-green mixture scale. Subjects G and H both passed the 'dichromatism' test and carried out the *normal* recipe. Subject H has a very high RGI value of 0.97, makes only 3 mistakes on the Ishihara, 1 mistake on the AO-HRR and passes the D15 and CU tests. Subject H's RG threshold is 9 units and from the plot of his results it is observed that his measured data points lie relatively close to a typical normal trichromat (i.e. the black ellipse shown in the figure).

The YB chromatic thresholds for all these observers do not show much variation.

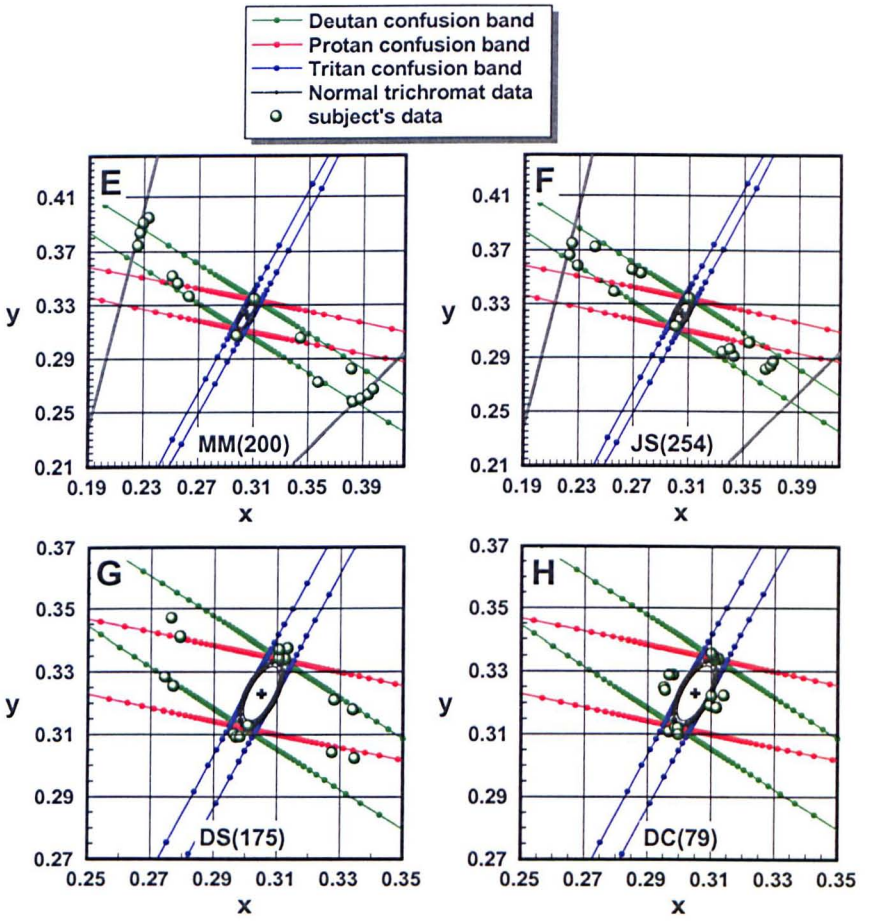


Figure 3.16: Chromatic threshold contours for four deutan observers plotted on the CIE-x,y chromaticity chart. See caption of Fig. 3.15 for details. The extent of colour vision loss is related to the elongation along the deuteranopic confusion band.

			CAD test		Colour Vision test					
Subject	M/F	Age	RG threshold	YB threshold	Nagel range	RGI	Ishihara	AO-HRR	D15†	CU
E	M	57	103	15	0-73	0	19	11	F	8
F	M	33	83	12	8-18	0.85	22	10	F	1
G	M	40	31	14	16-28	0.82	18	8	P*	0
H	M	33	9	13	17-18	0.97	3	1	P**	0

† F = fail, P = pass
* 2 adjacent transpositions
** 1 adjacent transposition

Table 3.4: Summary of results for subjects E-H in Fig. 3.16. For all four subjects the RG and YB threshold of chromatic sensitivity on the CAD test have been computed. See caption of Table 3.3 for explanation.

Subjects A-H, all show typical patterns of congenital colour vision deficiency. By measuring more angles, specifically located along the affected colour confusion axis, a better description of their chromatic sensitivity loss can be obtained.

3.3.2 Analysis of minimal colour deficient observers

The following subjects carried out the normal recipe (not the specific protan/deutan recipes), which tests for 16 directions in colour space: eight along the YB and eight along the RG direction with four directions each orientated for the protan and deutan axis. In order to distinguish protanomalous from deuteranomalous observers, we have looked at the symmetry of the thresholds measured in the RG direction. Fig. 3.15D shows clearly four of the measured directions that correspond more to the deutan colour confusion axis have smaller thresholds by comparison with the remaining four directions that correspond to the protan colour confusion line. For convenience we will describe the mean of the first four (deutan line) and the remaining four (protan line) as the green (G) and red (R) thresholds, respectively. Fig. 3.17 shows a plot of R versus G colour vision loss. The difference amongst these two types of deficiency is observed and is accentuated for higher chromatic thresholds. The difference occurs due to a difference in the YB signal in these two types of deficiency that extends the chromatic thresholds either along the particular protanopic or the deuteranopic confusion band. The observers with the lowest thresholds are usually deuteranomalous and the R and G thresholds are statistically indistinguishable. There is however one observer amongst these 78 included in Fig. 3.17, that is diagnosed as a protanomalous on the Nagel anomaloscope (range 56-65) but has symmetric thresholds on the CAD test, i.e. approximately the same loss of R and G thresholds.

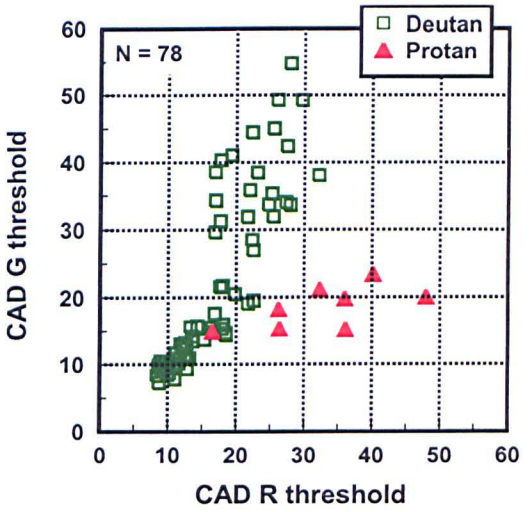


Figure 3.17: Results for 78 colour deficient observers tested on both the CAD test and the Nagel anomaloscope. Chromatic thresholds for the protanopic (R) and deuteranopic (D) axis are shown. For protanomalous observers the 'R' threshold is higher compared to the 'G' threshold. The opposite occurs for deuteranomalous observers. However for the milder cases of colour deficiency, with lowest chromatic thresholds, the thresholds for both 'R' and 'G' are statistically indistinguishable.

3.3.3 Repeatability in the measurement

The assessment of reliability of the CAD test provides some indication of the repeatability in the measurement. To make sure the CAD test is valid it must produce results that are repeatable. 33 observers were measured twice on the CAD test and the chromatic thresholds were computed as described in Section 3.2.6. Fig. 3.18 shows the results of the two measurements. The RG threshold data show a strong correlation (R^2 of 0.98), compared to the YB thresholds (R^2 of 0.81). The slightly higher correlation of RG thresholds may well reflect the higher noise level associated with YB discrimination. The data for the RG and YB thresholds includes both colour normal and deficient subjects explaining the spread up to 100 for the RG and 35 for the YB in CIE-x,y units. The spread suggests that the difference in repeated measures is greater for colour deficient observers, which is often the case, that variability is larger when the mean is larger. The variability observed in colour deficient observers is not clinically relevant from a diagnosis point of view. The outcome of abnormal or normal colour vision for an observer

will be the same in repeated measures, i.e. the variability is small to allow separating normal from colour defective observers. Fig. 3.19 is a plot of the same data such that sampling effects are mitigated.

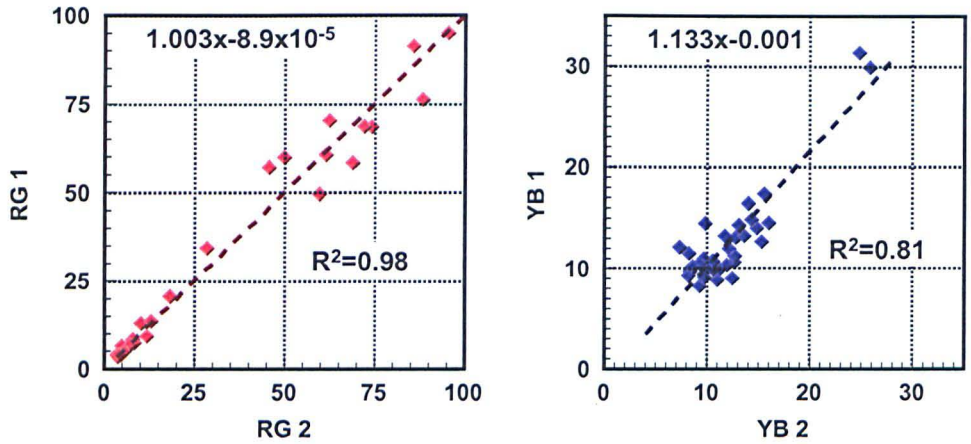


Figure 3.18: Scatter diagram of two measurements of chromatic thresholds performed on the CAD test for 33 observers. The comparison of the RG and YB thresholds shows that the data fits quite near to the line of unity slope. Best-fit lines and the correlation coefficients are shown.

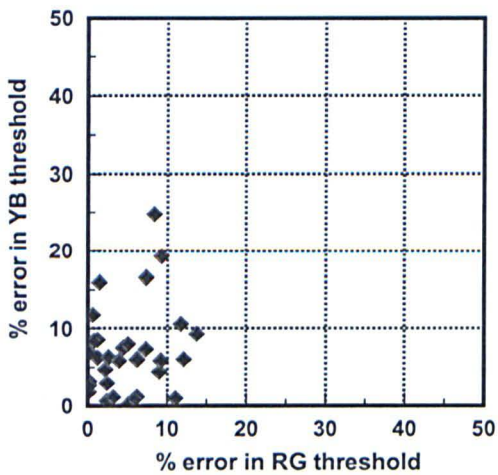


Figure 3.19: Plot of percentage error in a measurement for RG and YB threshold of chromatic sensitivity. The % error was taken as the absolute difference between the two repeated measurements divided by the sum of these two measures.

Fig. 3.19 is a slightly different visualization of the same data of Fig. 3.18 which shows how much the two measurements differ. On average, the subjects repeat the same result with an accuracy in the RG threshold measurement of ±5.0% (SD=0.7) and ±7.0% (SD=1.0) in the YB measurement. This plot suggests the percentage error is greater in the YB threshold compared to the RG threshold measurements also reflecting the increased noise in YB discrimination.

3.4 Discussion and conclusions

Advances in the use of colour displays, both in terms of increased applications and in the specific use of colour for coding information, highlights the need for more advanced methods of assessing colour vision. As result of the more extensive use of colour information in occupational environments, it has become important to understand how colour signals can enhance visual performance and the handicaps brought about by various levels of colour deficiency. The improvement and stability of high-resolution colour displays has provided an alternative medium to printed pseudoisochromatic plates that allows the measurement of chromatic sensitivity. A variety of test patterns can be configured and altered rapidly and the visual task is simplified so that a verbal response is not required.

Testing for YB chromatic sensitivity is an issue worth discussing. The diagnosis of tritan defects is difficult and the condition often eludes detection. For example some special pseudoisochromatic plates have been designed e.g. AO-HRR or by examining chromatic discrimination on pigment arrangement tests, such as the F-M 100 hue test or the D15 (both have a tritan axis). However, none of these tests alone or in combination, identify a tritan disturbance unequivocally. The CAD test, presented on a CRT display has all the advantages of selecting the colour directions one wishes to examine and therefore measuring tritan deficiencies as accurately as any type of congenital deficiency.

The results demonstrate clearly that the technique of spatiotemporal perturbation provides a means for effective isolation of chromatic signals. The CAD test makes

use of this technique and has been optimised to assess chromatic sensitivity. The effects of background luminance and stimulus size have been studied in order to measure chromatic sensitivity under optimum conditions. The CAD test quantifies the severity of colour vision loss by evaluating RG and YB colour detection thresholds using an internationally recognised colour system. The test is simple to administer: screening for severe colour deficiencies can be done in 90 seconds, and chromatic discrimination sensitivity can be obtained in 12 minutes from a naïve subject. The CAD test achieves good separation between severe, as well as among minimal protan and deutan colour vision deficient observers. The high correlation on re-testing is an indication of the small, overall noise levels involved. These reflect both the properties of fundamental RG and YB chromatic mechanisms and the accuracy of the CAD technique.

The CAD test provides a rapid means of screening subjects for colour vision deficiencies, but it also can be used to examine in more detail the changes in colour discrimination that occur as a result of congenital or acquired conditions (Barbur et al. 1994; Barbur et al. 1997). It allows the investigator to monitor quantitatively over time the progression or remission of disease. Many drugs affect chromatic sensitivity and the pharmacologist may find the test suitable for progress of disease or the outcome of treatment.

The usefulness of the CAD test would be greatly enhanced by providing a statistical description of the limits of chromatic sensitivity expected within normal trichromats. Such data would make it possible to establish whether small changes in chromatic sensitivity fall outside the normal range.

4 Towards Establishing Limits of 'Normal' Chromatic Sensitivity in Human Vision

Approximately 95% of humans exhibit 'normal' vision. Although these are all trichromats, it is evident that individual differences in pre-receptoral and receptoral absorption properties will result in individual variations in relative spectral sensitivity, and hence, chromatic discrimination, colour matching and colour perception. The purpose of this chapter is to investigate the variability of normal colour vision to establish more closely whether minimum colour deficient observers can be separated reliably from 'normal' subjects by measuring their chromatic sensitivity. By examining a large number of subjects on a number of different tests including the Nagel anomaloscope, we were also able to examine the efficiency of the new CAD test in relation to several currently used tests.

4.1 Introduction

It is well established that the severity of colour vision loss varies significantly amongst colour deficient observers (Alpern and Pugh, Jr. 1977; Alpern 1979). The variation in chromatic sensitivity in normal trichromats is less well documented and can be attributed to a number of different factors. Humans with normal trichromatic colour vision possess three distinct classes of cone photoreceptors in the eye. These contain S (short), M (medium) and L (long)-wave sensitive photopigments (see Fig. 1.11). Small genetic mutations or altered expression in L and/or M cone genes can cause changes in chromatic sensitivity. Other factors that may or may not be genetically related such as changes in the optical density of cone photoreceptors and/or variation in post-receptoral amplification of cone signals (see Chapter 5) can also cause significant changes in chromatic sensitivity (Barbur 2003).

Conventional colour vision tests are appropriate for detection of congenital colour vision defects, but are less useful when one wishes to study the small differences in chromatic sensitivity that exist amongst normal colour vision observers. Furthermore, most of these tests are qualitative (e.g. Ishihara plates, AO-HRR plates, D15) or at most semi-quantitative (F-M 100 hue). More recently the use of stable colour monitors and efficient visual psychophysical procedures for isolating the use of colour signals and for measuring colour detection thresholds has been extended to applications that involve clinical colour vision testing. The obvious handicap that has emerged from the use of such test is the lack of accurate description of variability within the normal population against which one needs to evaluate the more accurate and reliable results of novel tests.

The first requirement of a good colour vision test is to isolate the use of colour signals, such that the object and its background are isoluminant, only colour changes across the borders can be used to distinguish the object, with no luminance contrast cues. True isoluminance is, however, difficult to arrange in practice, without resorting to subject specific minimally distinct border or heterochromatic flicker photometry methods, which are too time consuming for convenient clinical assessment. For example, pseudoisochromatic plates use

patches of different reflectances that attempt to mask any luminance changes to reduce the likelihood of a test number being defined and seen by detecting luminance differences rather than colour differences. The CAD test, when optimised to isolate chromatic signals, is also based on luminance masking techniques, but in addition it employs dynamic LC masking, a technique that stimulates selectively the magnocellular pathways that exhibit high luminance contrast sensitivity.

The aim of this study was to investigate the variability within normal trichromats and hence to establish the 'Standard Normal Observer' (SNO) limits for the CAD test. In order to determine what can be considered to be normal performance, a large population sample was investigated. This is a relatively common method to determine reference intervals for any kind of variables in population studies involving a random sample of subjects, where no prior information about the prevalence of a particular condition studied is known (Harris and Boyd 1995). This technique was used to avoid any adverse diagnoses with the standard conventional colour vision tests. It is self-evident that the size of the population sample and the accuracy of the measurement techniques are critical in determining the outcome of the study. With the standard normal CAD observer limits established, we were then able to validate and compare the CAD test with the standard reference test, the Nagel anomaloscope, and several other conventional colour vision tests. Data from both normal trichromats and colour deficient observers have been examined in order to assess the efficiency (i.e. sensitivity and specificity) of the CAD test.

4.2 Methods and Subjects

Chromatic sensitivity was assessed using colour-defined moving stimuli buried in a background of random dynamic LC noise. Chromatic thresholds were measured along 16 directions in colour space and the subject's task was to report the direction of motion of the colour-defined stimulus. An efficient, four-AFC procedure is used to measure subject's chromatic thresholds by interleaving the

number of carefully selected directions in the 1931 CIE-x,y chromaticity chart. The colour directions are grouped together so as to test RG and YB colour sensitivity. Chromatic thresholds are computed by averaging eight directions each for both RG and YB chromatic channels. A full description of the CAD test is given in Chapter 3.

Other tests for assessing colour vision were employed including the Nagel anomaloscope, Ishihara plates, AO-HRR plates, D15 and CU tests. Test conditions were as described in Chapter 2.

4.2.1 Subjects

A total of 472 subjects were tested on the CAD test, consisting of three groups of subjects, the first a random population sample of 245 subjects that were recruited from around the university. These subjects were invited to participate in a study with no prior information on the status of their colour vision. 117 were female and 128 were male with ages ranging from 14 to 60 years of age. The age distribution of the sample of random subjects had a median of 26 years and is shown in Fig. 4.1. The age limit was set at 60 years to minimise the effects of advancing age. They were questioned about their health and if no problems were reported they were accepted for participation. This random sample of subjects is therefore assumed to be an accurate representation of the general population. Additionally, 22 subjects above 60 years and with no age limit were also tested (14 male, 8 female; age range 61 to 75) to investigate the extent to which age can affect colour vision. The third group were subjects tested at City University who had failed a previous colour vision examination elsewhere and needed a full assessment of their colour vision using a battery of colour vision tests or who already suspected that they had some sort of colour deficiency. These subjects were used to assess the extent to which one can rely on the results of the CAD test to separate 'normal' and colour deficient observers. Amongst this group of colour vision deficient patients 200 male, 5 female were examined, age range 7 to 57.

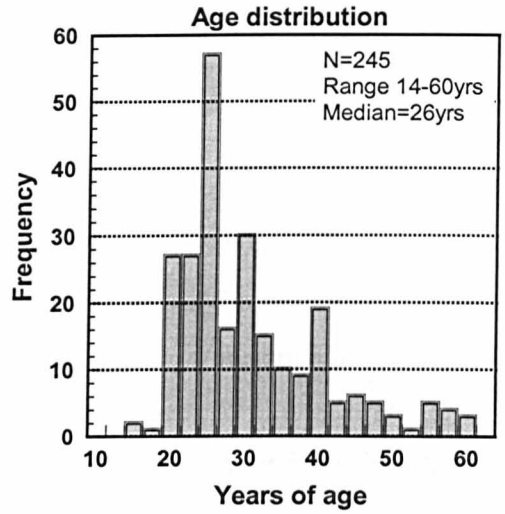


Figure 4.1: Age distribution of subjects from the random population sample tested in this study. The ages range from 14 to 60 years old with a mean of 29.7, standard deviation of 9.8 and a median of 26 years.

Out of the total of 472 subjects in this study, the following subject numbers performed the conventional colour vision tests: 231 subjects were measured on the Nagel anomaloscope, 311 on the Ishihara pseudoisochromatic plates, 242 on the AO-HRR, 240 on the D15 and 226 on the CU tests (see Table 2.2).

4.2.2 Statistical analysis

Nonparametric methods were employed to examine the data. These methods require no assumptions concerning the distribution from which the sample is drawn or any specific values of any parameters of that distribution. The methods employed to carry out comparisons between different colour vision tests involve similar procedures to those described in Chapter 2: including 2x2 frequency tables, the parameters of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and agreement (κ).

4.3 Results

4.3.1 Study of variability within normal trichromats

Initially, for identifying subjects with normal trichromatic vision we examined only the data from the random population sample. These data were treated independently of the results obtained in other colour vision tests. A scatter plot of RG versus YB chromatic thresholds measured in the 245 subjects is shown in Fig. 4.2. The distribution of points of the 245 observers (black squares) reveals one tight main cluster and eight outliers that are likely to be colour deficient observers with increased RG thresholds. It is worth mentioning that two of these 'outliers' (with RG thresholds of 12.6 and 13.5 CIE-x,y space units) were not aware of any sort of colour vision deficiency. They had never been examined before, as colour vision had never been a professional requirement. Amongst these eight subjects, one is female and seven are male. Therefore the resulting incidence of colour vision deficiency observed within this population sample is 5.5% for males and 0.8% for females. The prevalence of red-green colour deficiency found in large population surveys are of the order of 8% in men and 0.4% in women (see Section 1.6.2.1).

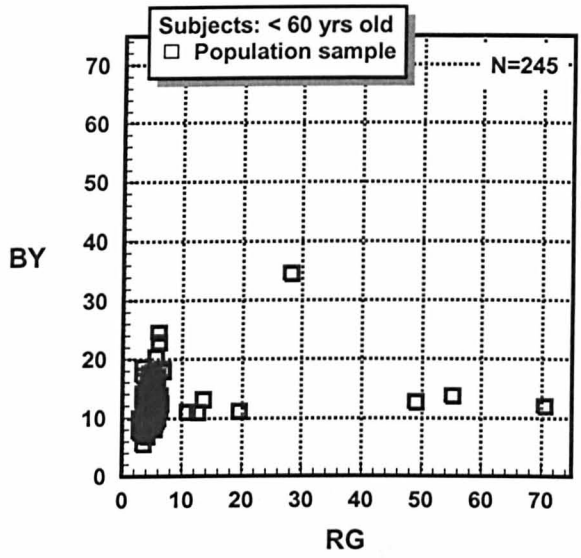


Figure 4.2: A scatter diagram of measured RG versus YB thresholds (1931 CIE-x,y space $x10^3$) for 245 observers, including 128 male and 117 female. The plot shows one main cluster that is clearly separated from 8 outliers which have higher RG thresholds.

Fig. 4.2 also reveals that among the main cluster of normal colour vision subjects, the spread in the RG thresholds is much less than the spread observed in the YB thresholds, not including the eight colour vision deficient observers that are clearly separated. This suggests that the normal limits on the CAD test based on chromatic thresholds are likely to exhibit smaller variance for RG than YB thresholds. The spread in the latter varies from about 5 to 25 units in 1931 CIE-x,y colour space. Amongst the eight outliers, four have a relatively moderate loss of RG sensitivity whilst four have a more severe loss and one subject shows significant loss both in RG and YB discrimination.

A brief analysis suggests that all those subjects with RG chromatic thresholds less than 7 units in the CIE-x,y colour space have normal colour vision and those eight outliers with RG thresholds above 10 have a RG colour vision deficiency.

In order to investigate whether this separation between the main cluster of normal trichromats and colour vision deficient observers still holds after inclusion of the group of presumably colour vision deficient individuals, the chromatic thresholds for the remaining subjects (less than 60 years old) were superimposed as shown in

Fig. 4.3. Fig. 4.3 shows the large spread in variability in RG thresholds along the abscissa. The inset shows clearly the gap between those observers that form the main cluster, identified in Fig. 4.2, and observers with minimal colour vision deficiency. The cluster on the left of the gap, the normal trichromats, shows low variability in RG thresholds when compared to the spread in RG thresholds of the remaining subjects. The data also suggest that a cut-off, or upper limit, of 7 RG thresholds units would easily distinguish 'normals' from (RG) colour deficient observers.

No exact cut-off limit is set for the YB threshold. It is obvious that there is one tritanopic observer and it is neither possible with so few tritanopic subjects nor necessary to define an exact limit at this stage. The total spread along on the YB axis, when including colour deficient observers, is relatively small in comparison to the RG thresholds.

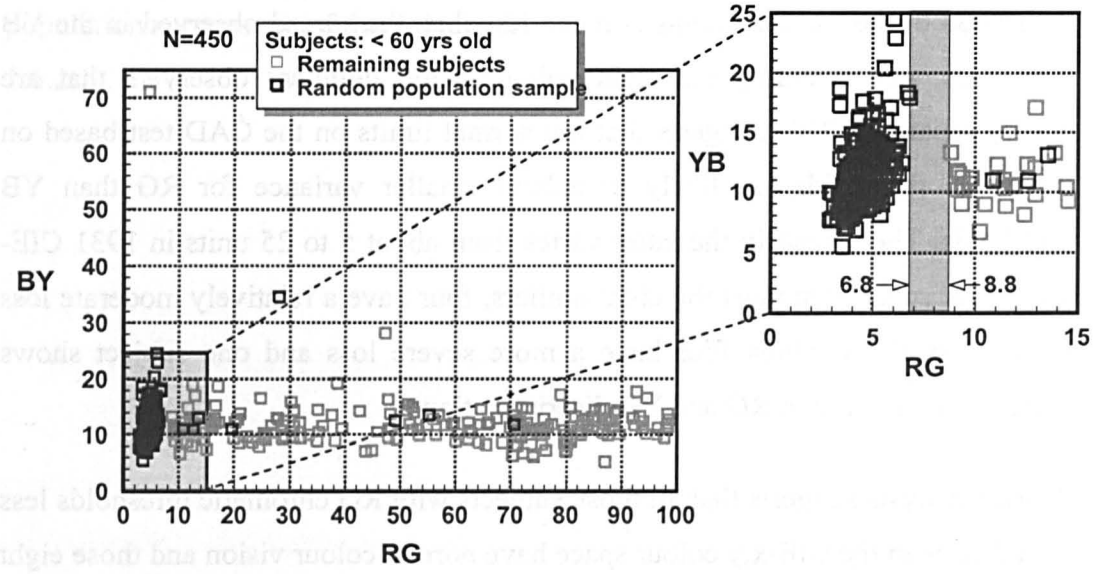


Figure 4.3: Plot of RG and YB thresholds (1931 CIE-x,y space x10^3) for the population sample shown in Fig. 4.2 (N=245, black squares) and the remaining subjects (presumably colour deficient, grey squares) making a total of 450 observers. On the plot on the left, a cluster of black squares can be observed concentrated at low RG thresholds. The data for deutan and protan colour deficient observers is spread along the abscissa illustrating clearly the spread in RG deficiencies that exists. The inset on the right shows the "gap" in RG thresholds that separates 'normal trichromats' from colour deficient observers.

The larger thresholds that affect both RG and YB discrimination in one of the subjects investigated are indicative of acquired deficiency (see Fig. 4.3).

For this data sample, setting a RG cut-off limit at 7 RG units in 1931 CIE-x,y colour space, allows to investigate the distribution and the variability within the 'normal' trichromats (N=237). Fig. 4.4 examines the distribution of the measured chromatic thresholds in turn for the YB and RG discrimination, and the ratio YB/RG for observers from the main cluster. One interesting observation worth drawing attention to is the difference in spread of RG thresholds by comparison to YB and the ratio of YB/RG thresholds. While the spread of YB and YB/RG follow approximately Gaussian distributions, the distribution of RG thresholds may indicate the existence of two or three overlapping subgroups, or at least a rather skewed one (discussed below). 95% of observers lie between the 2.5% and 97.5% empirical percentiles of the set of measurements. For the RG direction the 2.5^{th} percentile is 3.28 and the 97.5^{th} percentile is 6.10, and these values thus specify where 95% of the central range of observers lie. Similarly, the values 7.71 and 17.81 are the 2.5^{th} and 97.5^{th} percentiles respectively, of the observed distribution of YB thresholds. Due to the skewness of the data, this empirical estimation of percentiles is more appropriate to describe the variability within the YB and RG thresholds, than a parametric approach of taking the mean ± 1.96 standard deviations, which assumes a Gaussian distribution of the data.

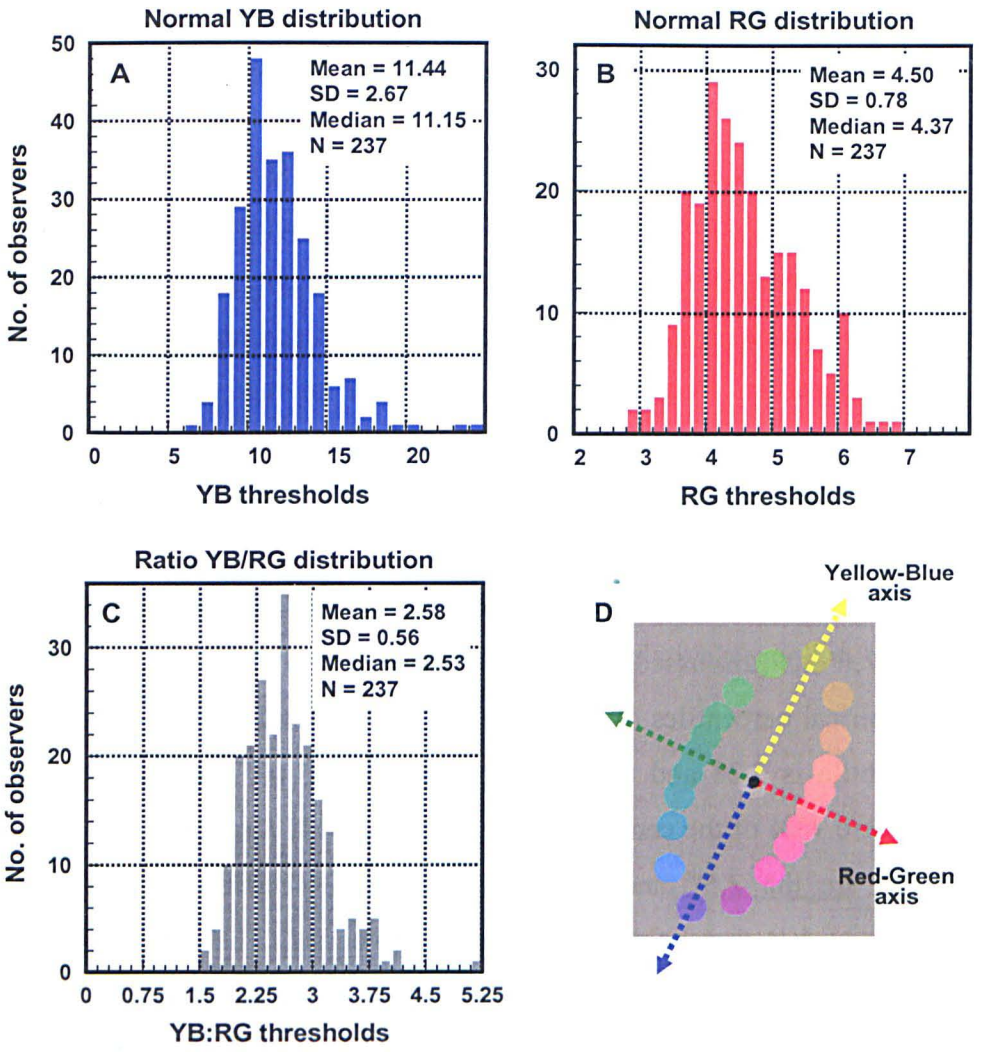


Figure 4.4: Frequency distributions of the following chromatic thresholds: (A) YB, (B) RG and (C) ratio YB/RG, obtained in 237 observers with 'normal' trichromatic vision. The YB and RG thresholds were computed by averaging 8 directions in colour space. The mean standard deviation (SD) and median are shown. (D) shows the two directions in 1931 CIE-x,y space used to compute both the YB and RG chromatic thresholds.

An interesting aspect worth of investigation are the differences amongst the male and female populations for the RG distribution. Fig. 4.5 shows the same data as in Fig. 4.4B, but this time separated according to sex.

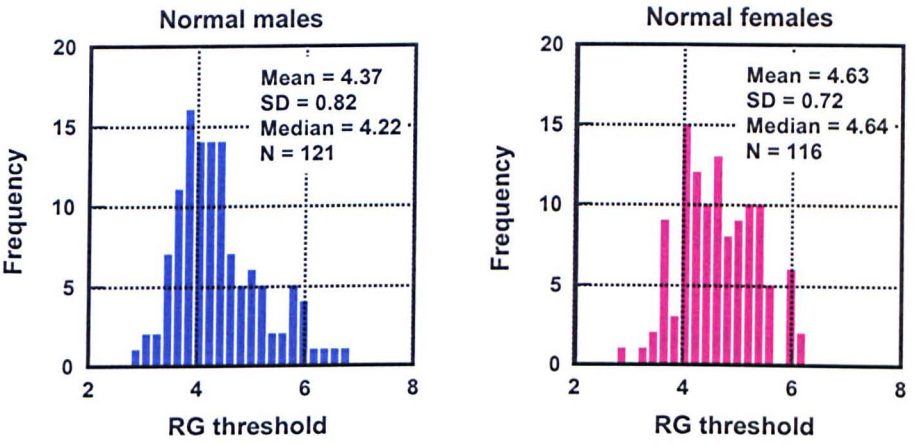


Figure 4.5: Histograms showing the different distributions in RG thresholds classified into male and female. Mann-Whitney test showed a significant difference in the medians between the two distributions (p=0.002).

The two distributions plotted for males and females are different (p=0.002). The differences in these two patterns suggest that the mechanisms underlying RG colour vision are inherited as an X-chromosome linked trait (see Chapter 1 Section 1.6.2.1). The data shows that females on average tend to have higher RG thresholds compared to males. The distribution in males might give indications of one large subgroup with thresholds around 4 and one or more smaller subgroups with lower sensitivity (i.e. with higher thresholds). An explanation for this difference suggested by Neitz (Neitz and Jacobs 1986; Neitz et al. 1995) is a polymorphism in the L-cone amongst males with normal colour vision. Amongst females, no such separation into two subgroups seems to exist. If this polymorphism alters the RG sensitivity, there would be three genotypic classes in females (compared to two in males), and the large group of heterozygotes leads to such an overlap in sensitivity that the three groups cannot be distinguished clearly.

Fig. 4.6 shows a graphical representation in the 1931 CIE-x,y colour chart of the statistical limits of chromatic thresholds amongst normal trichromats measured in this study. The percentiles calculated above allows us to plot ellipses in the 1931 CIE-x,y colour space that show the range within which the central 95% of values lie (i.e. excluding 2.5% at each end of the distribution). The 2.5th percentile is defined by the lower limit of the YB threshold (major axis of ellipse) and the

lower limit of the RG threshold (minor axis of ellipse). Similarly, the 97.5th percentile is defined by the upper limit of the YB threshold (major axis of ellipse) and the upper limit of the RG threshold (minor axis of ellipse). The ellipse orientation for normal trichromats was 62 ± 7°, defined by calculating the average direction with the highest YB thresholds for all subjects. The grey shaded area shown represents the region of the CIE chart where 95% of the observers' chromatic thresholds are expected to lie. The 2.5% and 97.5% limits represent the boundaries of this area. The median chromatic discrimination threshold ellipse is plotted. It is worth pointing out that the data show some evidence of skewness but that this does not affect the method used to compute the percentiles. The median ellipse is thus closer to the 2.5% limit than the 97.5%.

Fig. 4.6 is an extremely useful representation in that it provides a template defining the SNO on the CAD test in which a subject's individual measurements can be plotted allowing an instant diagnosis of either normal or deficient colour vision. If the observer's data points fall inside the grey shaded area, one can be highly confident of normal colour vision. In some situations, several of the data points might fall close to the 97.5% limit or slightly outside of this contour. If an observer's data points fall outside the 97.5th percentile contour, this suggests deficiency and the other way around, if an observer's data points fall on the lower side of the 2.5th contour, this suggests better than normal chromatic sensitivity. In circumstances where average thresholds are greater than the 97.5th percentile along the RG direction, the mean and its standard deviation is calculated. If the difference between the mean threshold and the expected 97.5% limit for the SNO is less than two standard deviations, the subject is considered as having normal colour vision. If this is not the case, the obvious conclusion is that the subject's performance cannot be included within the range that describes 97.5% of the normal population.

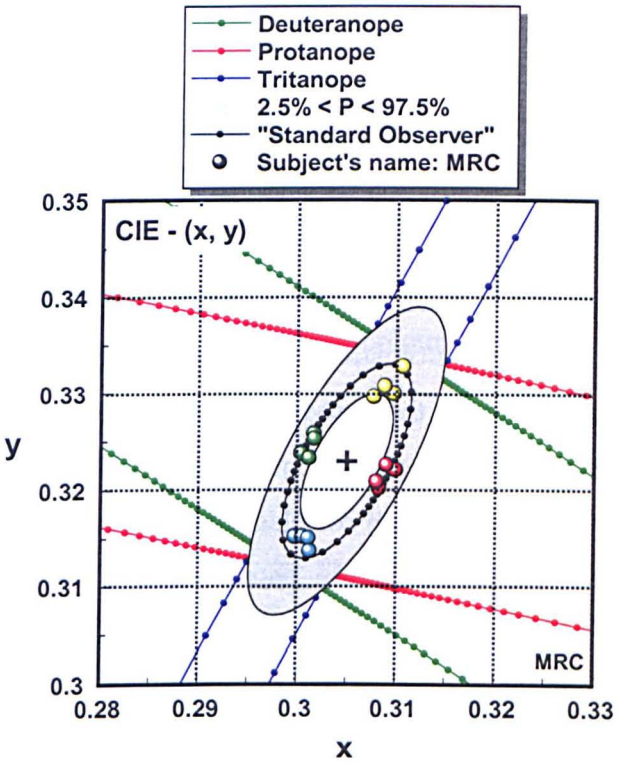


Figure 4.6: Data showing the 97.5 and 2.5% statistical limits that define the "standard" normal CAD test observer. The dotted, black ellipse is based on the median RG and YB thresholds measured in 237 observers. The grey shaded area shows the limits of variability of 95% of these observers. The deuteranopic, protanopic and tritanopic confusion bands are displayed in green, red and blue, respectively. The background chromaticity (x,y) is indicated by the black cross (0.305, 0.323). The coloured symbols show data measured for a typical normal trichromat (RG=4.23 ± 0.16; YB=9.08 ± 0.42).

This study of the variability within normal trichromatism allows defining the SNO in terms of CAD units, where the major axis of the median ellipse is 1 YB CAD unit and the minor axis 1 RG CAD unit. The median threshold for RG is 4.37 and for YB is 11.15 in CIE-x,y colour space (see Fig. 4.4). Using this definition of the SNO the distinction between normal and defective colour vision and a quantification of colour vision loss can be expressed in a way that is simple and easy to understand.

In summary, this leads to the following definitions: a RG threshold of less than 3.28 implies better than normal colour vision, a threshold between 3.28 and 6.1 is considered to be within the normal range, subjects with a threshold between 6.1 and 7 cannot be included within the 97.5% range of the normal population, but

may have normal colour vision, and a threshold above 7 leads to a diagnosis of colour vision deficiency.

In addition, the effect of age on chromatic thresholds for the population sample of 237 subjects (having eliminated the 8 outliers) can be studied. Fig. 4.7 shows a plot of chromatic thresholds versus age for the random population sample. Subjects above age 60 years have been added to reveal the extent to which advancing age may affect chromatic discrimination sensitivity, although their thresholds do not contribute to the best-fit line. These results show that up to the age of 60 there is little effect on the chromatic thresholds. There is a small effect on YB thresholds (but no significant correlation) and an even smaller effect of age on RG thresholds. This supports the choice of age ranges for the random population sample in this study. Our choice seems appropriate since most tests are designed to assess colour vision deficiencies for occupational needs. This happens during the typical working life, usually at the beginning of one's career and sometimes periodically afterwards.

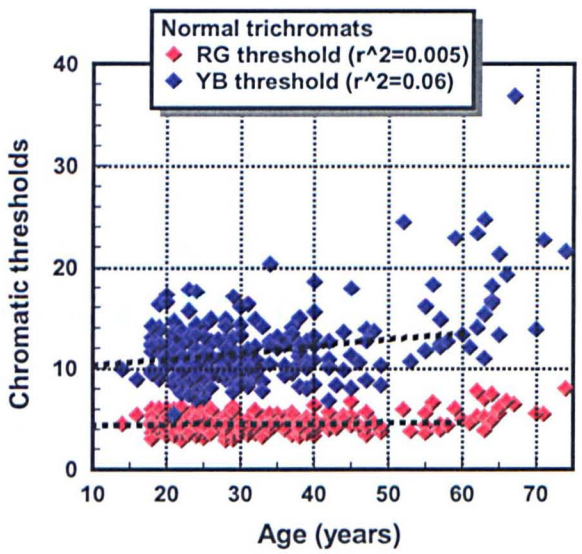


Figure 4.7: Effect of age on YB and RG chromatic thresholds. The best-fit line is shown for both sets of data. The correlation coefficients (r^2) for the RG and YB thresholds are 0.005 and 0.06, respectively.

4.3.2 Validation of the CAD test against the Nagel anomaloscope

The validation is done with all 450 subjects.

In the previous section the definition of the SNO on the CAD test makes it possible to re-plot Fig. 4.3 in normal CAD units for YB and RG thresholds. Fig. 4.8 shows the data plotted so as to identify the observer's type of colour vision deficiency (diagnosed using conventional colour vision tests as described in Chapter 2).

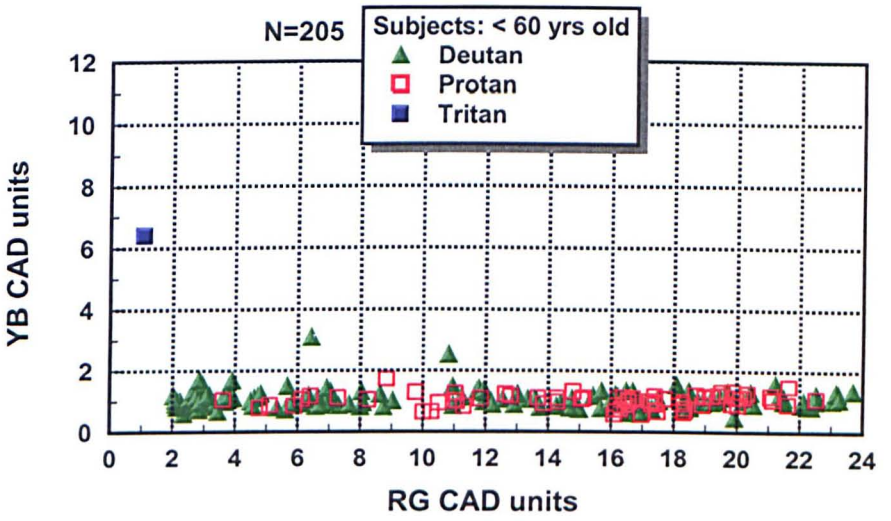


Figure 4.8: Plot of RG and YB thresholds in CAD normal units for the remaining colour vision deficient subjects in Fig. 4.3 (N=205). The data for deutan and protan colour deficient observers (represented by triangles and outlined-square, respectively) spread along the abscissa from twice the normal threshold units. The only tritan observer tested, represented by a filled square, reveals a loss of over 6 times the normal CAD units.

Fig. 4.8 also shows that the mildest deficiencies are usually those by deutan colour deficient observers, as low as twice the 'normal' CAD thresholds. Protans lowest thresholds, on the other hand, are around four times the SNO and their concentration is highest at thresholds greater than 16 times the SNO. A point to note here is the expectation of there being many more colour deficient observers with lower RG thresholds, especially amongst protan colour deficient observers. An argument for this lack of mild deficiencies is that these people might not be

seriously impaired by their colour vision loss so unaware of the condition and undetected by standard colour vision tests. Such subjects may have not been referred to us for an assessment using the CAD test.

In order to validate the CAD test it must be compared to a previously established colour vision test. Other tests such as the Ishihara pseudoisochromatic plates, AO-HRR plates, D15 and CU tests have been validated against the Nagel anomaloscope in earlier studies (see Chapter 2).

To assess the efficiency of the CAD test in comparison to the Nagel anomaloscope, the definition of the normal observer on both tests is used. The normal observer on the standard reference test (Nagel Anomaloscope) was determined by examining 231 subjects and, from these, 70 formed a distinct cluster that was clearly separated from all others. The red-green colour matches measured (midpoints and matching ranges) for these 70 subjects were used to calculate the limits that define the 'normal' trichromat on the Nagel anomaloscope (see Chapter 2). Based on this analysis, the 'normal' Nagel parameters are defined as follows:

- The midpoint of the colour match should be within 36-44 on the red-green mixture scale and;
- the range of the colour match should be less than seven units on the red-green mixture scale (see Section 2.3.1.3 for a description of this definition).

Reverting back to the CAD test, the criteria employed to separate colour normals from colour deficients is the definition of the 'normal' RG threshold limit in 1931 CIE-x,y colour space. It is important to note that the YB CAD data cannot be compared with the Nagel as the latter only tests for red/green colour deficiencies.

4.3.2.1 CAD test versus Nagel anomaloscope

Table 4.1 compares the results of the two colour vision tests against each other with well defined standard 'normal observers' on each test. A pass in one colour vision test implies satisfying the criteria of the normal colour vision observer on that test. If the criteria are not met, the test is failed and colour deficiency is implied under the current test.

		Nagel Anomaloscope Midpoint 36-44 Range 1-7		
		Pass	Fail	
CAD test SNO RG axis only	Pass	63	4	κ=0.97 Specificity=100% (95% CI: 0.93-1)
	Fail	0	156	Sensitivity=97.5% (95% CI: 0.93-0.99) 223 subjects

Table 4.1: Contingency table showing number of subjects that pass/fail the normal observer criteria for both the CAD test and the Nagel anomaloscope. 223 observers were examined on both tests and the comparison reveals an agreement of 0.97 (where 1 signifies perfect agreement). The efficiency of the CAD test versus the Nagel anomaloscope shows a specificity value of 100% and a sensitivity value of 97.5%.

A total of 223 subjects (excluding those >60 years old) carried out both the Nagel anomaloscope and the CAD test. Table 4.1 shows that 63 passed both tests, 156 failed both tests and four subjects failed the Nagel but passed the CAD test. The agreement (κ) of the CAD test compared to the Nagel is 0.97, where $\kappa = 1$ indicates perfect agreement and zero indicates that agreement is at the level of chance. The specificity of the CAD test compared to the Nagel is 100%, in other words 63 subjects with red-green matches within the statistical limits of the normal observer on the Nagel anomaloscope also satisfy the SNO of the CAD test. When dealing with anomalous colour vision, the sensitivity of the CAD test is 97.5%; four subjects fail the Nagel anomaloscope criterion for normal observer, however are within the 'normal' CAD threshold units. Table 4.2 shows the results

of these four subjects in more detail. These four subjects did not satisfy the normal observer on the Nagel anomaloscope either by failing the midpoint or the matching range condition. Subjects 119 and 144 failed to have midpoints within the range (i.e. 36 and 44), while subjects 116 and 218 failed to have a matching range less than seven on the red-green mixture scale (see Table 4.2). These four subjects however satisfy the SNO on the CAD test as shown by their plots in Fig. 4.9 and the RG thresholds (Table 4.2).

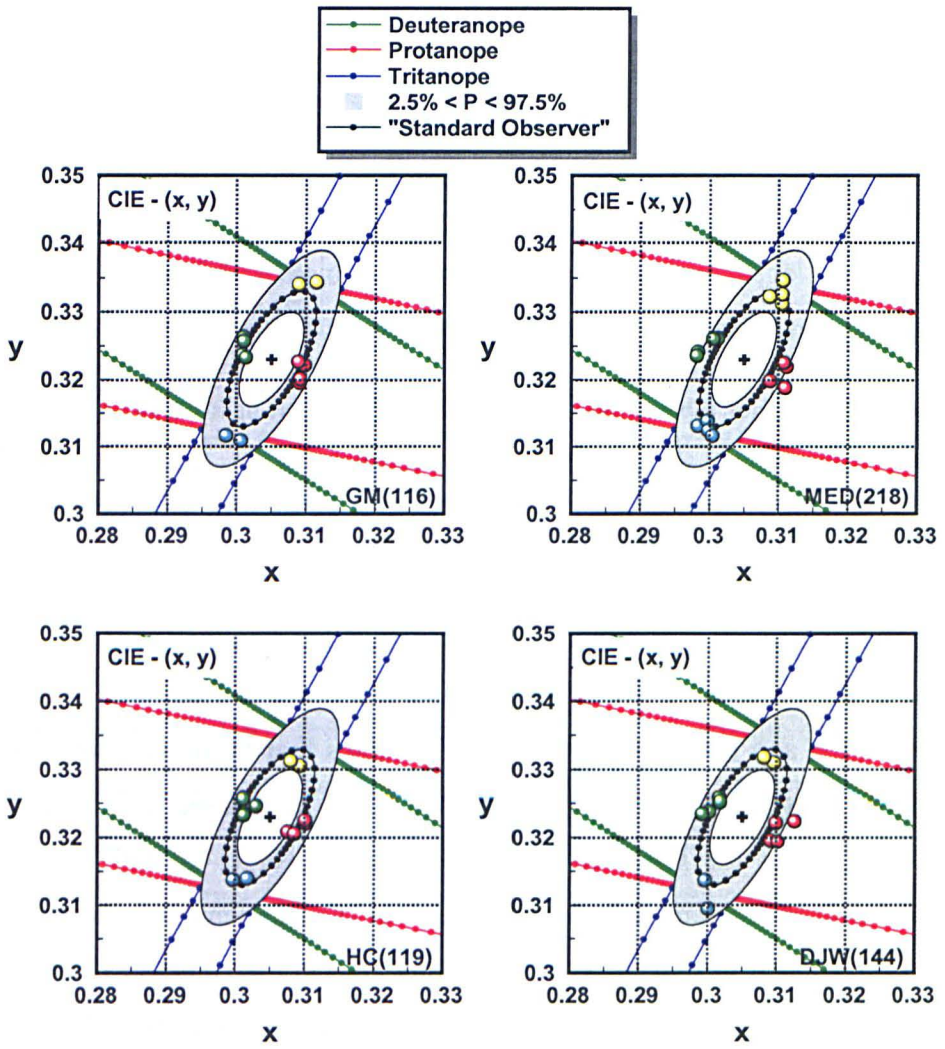


Figure 4.9: Plots of the chromatic thresholds for the subjects in Table 4.2. These four subjects are not within the 'normal' on the Nagel anomaloscope, however satisfy the SNO on the CAD test.

			CAD test		Nagel Anomaloscope			
Subject	M/F	Age	RG threshold	YB threshold	Match	Range	Midpoint	RGI
116	F	34	4.806	12.677	38-46	9	42	0.88
119	M	19	4.427	9.437	44-45	2	44.5	0.97
144	M	42	5.742	10.929	31-35	5	33	0.93
218	F	26	6.031	11.258	39-46	8	42.5	0.89

Table 4.2: Subjects that failed the normal on the Nagel and passed the SNO on the CAD test (i.e. RG threshold less than 6.1).

Subject number 144 could be an exception and classified as a deuteranomalous on the Nagel anomaloscope, and subject 119 could be considered to have normal colour vision and his Nagel result might be due to a random measurement error. However, the other two, subjects 116 and 218, cannot be classified into any category easily. Several studies in the past have called observers that make these sorts of colour matches as 'colour weak' or 'deviants' (Pokorny et al. 1979; Wyszecki & Stiles 1982). In order to find out whether these subjects have normal or deficient colour vision further assessment on different colour vision tests is necessary (such as on the CAD test).

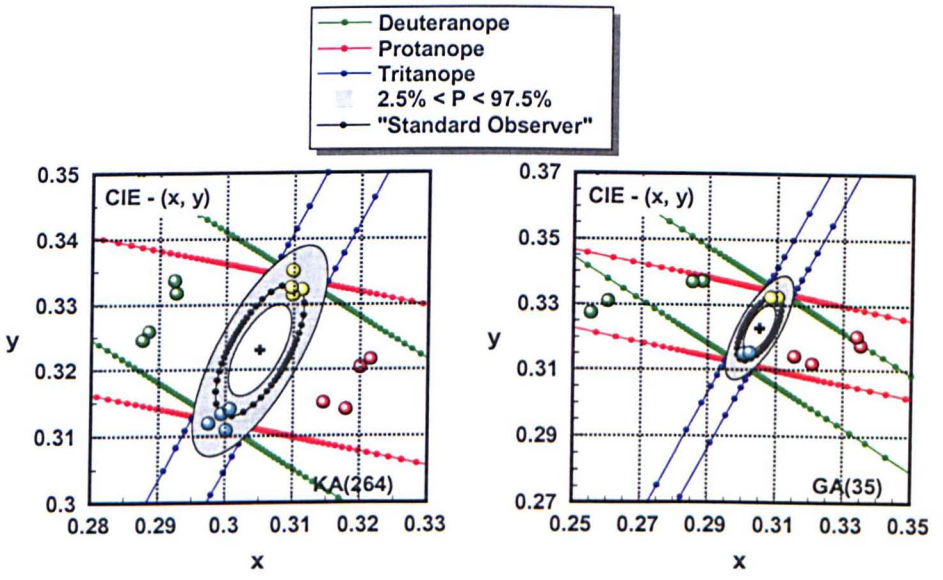
Different criteria for passing the Nagel, other than that based on the distribution of measurements taken from this population sample, could be set. If this was the case, the outcome of the comparison between the CAD and Nagel anomaloscope could be quite different. For example, the JAA (see Chapter 2) accepts a range of 4 units or less as criteria for passing the Nagel anomaloscope. Based on the JAA protocol the number of observers that pass the CAD test but fail the Nagel would increase to 18 for this population sample. Also the number of subjects that fail the CAD and pass the Nagel would increase to 25. The number of subjects that pass and fail on both tests would then be 38 and 142, respectively. Hence, the agreement would be lower, κ=0.64 and the specificity and sensitivity of the CAD test would be 60.3% (95% CI: 0.47-0.72) and 88.8% (95% CI: 0.83-0.93), respectively.

This analysis can be extended to examine the diagnosis efficiency of deutan or protan defects classified on the CAD and the Nagel anomaloscope. On the CAD test, the type of deficiency was classified by identifying the orientation of the highest RG thresholds; 140-155 and 165-180 for deutan and protan colour vision deficiencies, respectively (see Chapter 3). The Nagel anomaloscope classifies type of colour vision deficiency according to the proportion of green to red light the subject needs to match the yellow field compared to the normal match. Deutans make matches that contain a higher proportion of the green primary than the normal match, while protan matches contain a higher proportion of the red primary than the normal match (see Chapter 2). Normal matches therefore lie in between those of deutan and protan observers; these were identified on the scatter plot in Fig. 2.12 shown by the blue squares. This classification reveals any subject that is diagnosed on the anomaloscope and the CAD tests. Results are shown in Table 4.3; one tritan colour deficient has been added making the total number of subjects examined on both the CAD test and Nagel anomaloscope 224. Subjects 116, 119, 144 and 218 (see Table 4.2) have been labelled unclassified on the Nagel anomaloscope, (U) column in Table 4.3 (not fall into any category easily).

		Nagel Anomaloscope Midpoint 36-44 Range 1-7				
		D	P	N	U	κ =0.99
CAD test SNO	D	106	1	-	-	224 subjects
	P	1	48	-	-	Diagnosis; D (deutan) P (protan) N (normal) T (tritan) U (unclassified)
	N	-	-	63	4	
	T	-	-	1	-	

Table 4.3: Contingency table showing number of subjects classified according to diagnosis of type of colour vision on both the Nagel anomaloscope and CAD tests. The number of subjects diagnosed as having deutan colour deficiency (including deuteranomalous and deuteranopes) on the Nagel anomaloscope that are also classified as deutan on the CAD. Similarly, for protan colour deficiencies. One tritan colour deficient observer was revealed with the CAD test and made a 'normal' match with the Nagel anomaloscope. The four subjects, described in the text, which passed the CAD test and failed the 'normal' on the Nagel have been included here.

In conclusion, the CAD test shows excellent diagnostic efficiency of 0.99 with the 'gold-standard' test. Table 4.3, does however show three disagreements in classification between the CAD test and Nagel anomaloscope; two observers are misclassified as protan and deutan, respectively (see Fig. 4.10), and one observer performing normally on the Nagel but revealing tritan deficiency on the CAD test. The four unclassified observers have been previously described (Table 4.2): three could possibly be diagnosed as colour normal and one deutan on the Nagel anomaloscope. The tritan subject emphasises the fact that the Nagel cannot be used to test any YB colour vision loss; therefore further tests are needed to identify these types of deficiencies.



			CAD test		Nagel Anomaloscope			
Subject	M/F	Age	RG threshold	YB threshold	Match	Range	Midpoint	RGI
264	M	30	15.728	11.533	56-65	10	60.5	0.86
35	M	26	31.888	9.472	20-26	7	23	0.91

Figure 4.10: Plots of the chromatic thresholds and summarised results for two subjects misclassified in Table 4.3. The classifications on the Nagel anomaloscope and on the CAD test disagree for these two subjects. Subject no. 264 is classified as deutan on the CAD test, whilst on the Nagel his colour match is typical of a protanomalous observer. Subject no. 35 shows slightly higher thresholds along the protanopic axis on the CAD test whilst performs a typical deuteranomalous match on the anomaloscope.

If we now take the results from all subjects and compare the RG chromatic thresholds obtained on the CAD test to the Nagel anomaloscope RGI, we show that the correspondence is not exact. Fig. 4.11 shows the results for those subjects identified as colour deficient on the Nagel anomaloscope separated into protan and deutan deficiency. Here the striking feature is the large range of discrimination abilities exhibited by different subjects: RGI values range from 0 to close to 1 and 'normal' CAD RG threshold units go from about 2 to 24. In general, subjects with low discrimination abilities (RGI<0.4) have high RG thresholds (>15 with one exception). Subjects with higher RGI (>0.8) may also have lower RG thresholds. For deutans, the majority of these subjects have RG thresholds <10 and whilst for protans the RG thresholds are distributed rather evenly between about 4 and 20. Dichromats (RGI=0) always have large CAD thresholds of at least 15. Lower RG thresholds (<7) are only exhibited by subjects with higher RGI (>0.7). But besides these general patterns, the individual variation is very large. For example, there are several dichromats (RGI=0) who have a RG threshold on the CAD test of the order of 15, but there are also several protanomalous and deuteranomalous trichromatic subjects even with RGI>0.8 with RG thresholds of 15 units and higher. There are subjects with an RGI value close to 1 and their RG thresholds vary from 2 to 17 for deutan observers and from 5 to 17 for protan observers. These findings illustrate the little value the Nagel range or corresponding RGI has as a measure of chromatic sensitivity.

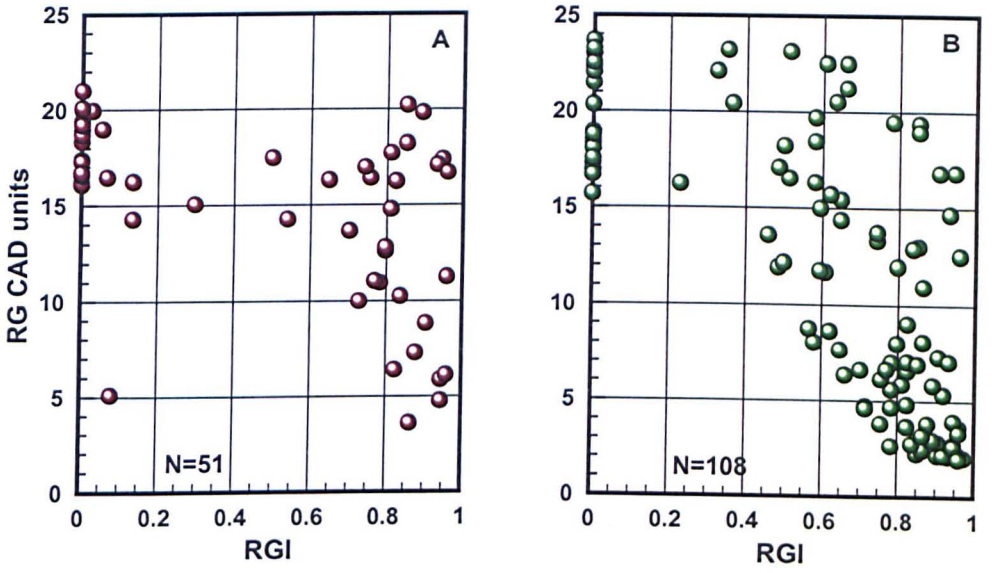


Figure 4.11: Results of colour deficient observers measured on the CAD test and Nagel anomaloscope. (A) 51 protan and (B) 108 deutan observers. The results show that the extent of the Nagel range (indicated by the RGI value) shows no strong correlation with chromatic sensitivity thresholds measured on the CAD test (measured in 'normal' CAD threshold units along the RG direction in CIE-x,y colour space). The correlation coefficient for best linear fits is (A) 0.27 and (B) 0.54.

4.3.2 Comparison of the CAD test against other colour vision tests

A comparison of several occupational colour vision tests to the CAD test can be done in a similar way to the previous analysis performed in Chapter 2 with the Nagel anomaloscope. Each test will be compared in turn with the CAD test and at the end there will be a section summarising all the tests.

4.3.3.1 Ishihara test

Out of 304 subjects tested on both CAD and Ishihara pseudoisochromatic plates (Table 4.4), both tests give the same diagnosis for 83 subjects that pass and 204 subjects that fail both tests. Only one observer passes the Ishihara plates and fails the CAD test. On the other hand 16 observers pass the CAD test and read the

Ishihara plates incorrectly. A reason for the low agreement of the Ishihara (among those failing the Ishihara) can be found by examining more closely those observers fulfilling the criteria of the SNO of the CAD test and failing the Ishihara plates. From the 16 normals on the CAD test that failed to pass the Ishihara plates, five subjects would have passed if up to two misreadings only and no errors are permitted. The errors of the remaining 11 subjects were mostly on the "hidden digit" plates (the ones without any number). Birch (1997c) recommends the use of the first 16 plates for screening for colour vision, which means excluding these "hidden digit" plates. If we now evaluate the Ishihara plates versus the CAD test again, but only for plates 1-16, only two or nine out of the 16 still fail the Ishihara test with and without allowing for misreadings, respectively. The best agreement of the Ishihara plate test versus the CAD test is when plates from 1-16 are used and up to two misreadings are allowed. For this case the agreement would be κ=0.98.

		CAD test SNO RG only		
		Pass	Fail	
Ishihara test				κ =0.91
Standard version 38-plate 1-25 plates (no errors & no misreadings)	Pass	83	1	304 subjects
	Fail	16	204	

Table 4.4: Contingency table showing number of subjects that pass/fail the Ishihara plate test and the normal observer criteria on the CAD test. 304 observers were examined on both tests and the comparison reveals an agreement of 0.91.

Similar results are shown when the number of errors on the Ishihara pseudoisochromatic plates is compared to the CAD RG thresholds. Fig. 4.12 suggests that the number of errors performed on this test by colour deficient observers does not necessarily correspond to a measure of chromatic sensitivity loss. In general, subjects with high RG thresholds make more errors (and vice

versa) but the individual variation is again large. For both protan and deutan subjects with more than 15 errors on the Ishihara plates, the RG chromatic thresholds may vary from 4 to about 23.

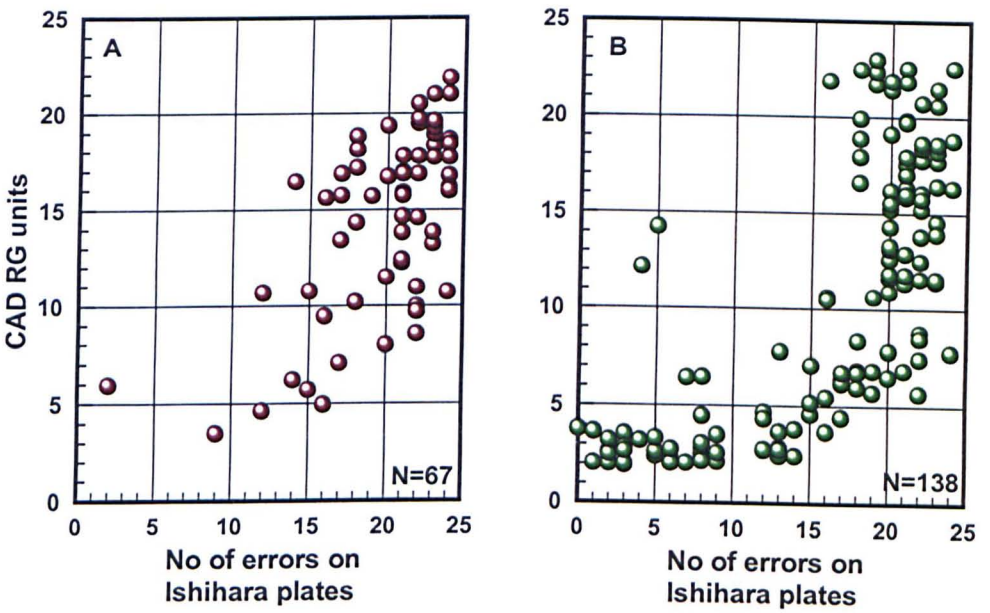


Figure 4.12: Results of colour deficient observers measured on the CAD test and Ishihara pseudoisochromatic 25 plates. (A) 67 protan and (B) 138 deutan observers. The results show that the number of failed Ishihara plates does not fully correspond to chromatic sensitivity thresholds measured on the CAD test. The correlation coefficient for best linear fits is (A) 0.40 and (B) 0.52.

4.3.3.2 AO-HRR test

A total of 235 observers carried out both the AO-HRR plate test and the CAD test. These results are shown in Table 4.5 below. From the 35 subjects that passed the CAD test, one failed the AO-HRR plate, making only 1 error (subject no. 218, results shown in Fig. 4.9 and Table 4.2). The 12 subjects that failed the CAD test and passed the AO-HRR, were all deuteranomalous observers and amongst these, five made Rayleigh matches of less than 7 scale units and four subjects made

matches between 7 and 13 units on the Nagel anomaloscope. 188 subjects failed both tests.

		CAD test SNO RG only		
		Pass	Fail	
AO-HRR (no errors)	Pass	34	12	κ =0.84 235 subjects
	Fail	1	188	

Table 4.5: Contingency table showing number of subjects that pass/fail the AO-HRR plate test and the normal observer criteria on the CAD test. 235 observers were examined on both tests and the comparison reveals an agreement of 0.84.

Fig. 4.13 shows the comparison between the number of errors on the AO-HRR plate test and the loss of chromatic sensitivity in 'normal' CAD units for the RG direction only. The results show that overall protan colour deficient observers make fewer errors on the AO-HRR test than deutan colour deficient observers. The number of failed AO-HRR plates shows some correlation with chromatic sensitivity thresholds measured on the CAD test.

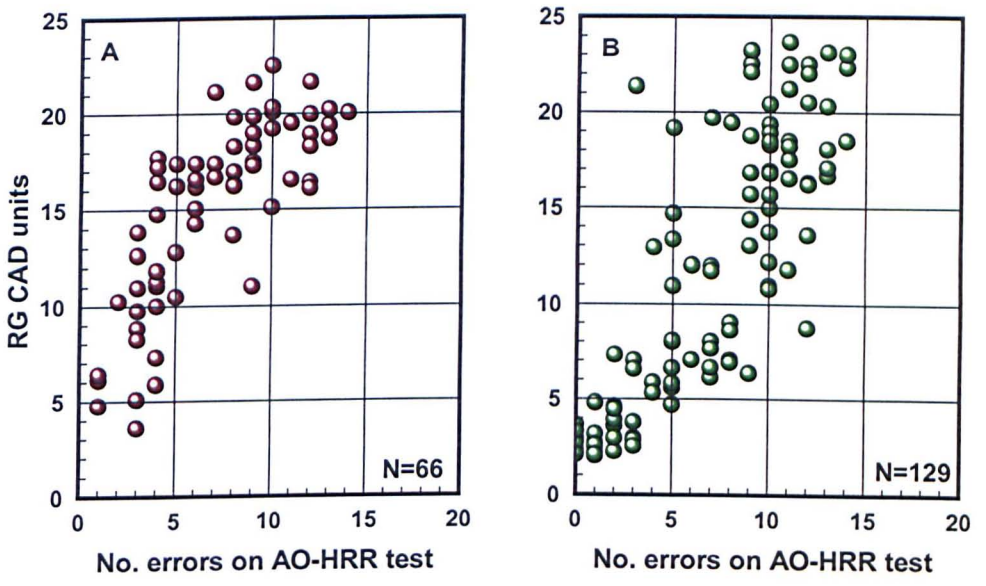


Figure 4.13: Results of colour deficient observers measured on the CAD test and AO-HRR pseudoisochromatic 20-plate test. (A) 66 protan and (B) 129 deutan observers. The results show a slight correlation between the number of failed AO-HRR plates versus RG chromatic sensitivity thresholds measured on the CAD test. The correlation coefficient for best linear fits is (A) 0.59 and (B) 0.71.

4.3.3.3 D15 and CU tests

A total of 234 subjects carried out both the D15 arrangement test and the CAD test. Table 4.7 shows that from the subjects that failed the CAD test, 95 passed the D15 when allowing a maximum of two adjacent transpositions, accounting for the low agreement with the D15 test (κ =0.39). From the 32 subjects that passed the CAD test, one failed the D15 test, making three adjacent transitions and one crossing between caps 1 and 5 (subject 218, see Fig. 4.9 and Table 4.2).

		CAD test SNO RG only: Pass	Fail	
D15 Max. of 2 adjacent transpositions	Pass	31	95	κ =0.39 234 subjects
	Fail	1	107	

Table 4.6: Contingency table showing number of subjects that pass/fail the D15 test and the normal observer criteria on the CAD test. 234 observers were examined on both tests and the comparison reveals an agreement of 0.39.

Similar results are obtained from comparing the CU test and the CAD test as shown in Table 4.7. Of those subjects that failed the CAD test 78 passed the CU test giving a low agreement of κ =0.42. From the 28 subjects that passed the CAD test, these also passed the CU test, thus giving a specificity of 100%.

		CAD test SNO RG only: Pass	Fail	
CU (no errors)	Pass	28	78	κ =0.42 220 subjects
	Fail	0	114	

Table 4.7: Contingency table showing number of subjects that pass/fail the City University test and the normal observer criteria on the CAD test. 220 observers were examined on both tests and the comparison reveals an agreement of 0.42.

4.3.3.4 Overall agreements with the Nagel anomaloscope

The comparison of all the colour vision tests employed in the previous and the current chapter, including the CAD test, versus the gold standard Nagel test as the reference is given in Table 4.8 below.

			Sample Prevalence City University			Population Prevalence Males ~0.08		Females ~0.004	
Test	Sens†	Spec††	Prev.	PPV[1]	NPV[2]	PPV[1]	NPV[2]	PPV[1]	NPV[2]
CAD	97.5	100	0.72	100	94	100	99.8	100	100
Ishihara	100	87.5	0.87	98	100	41.0	100	3.1	100
AO-HRR	94.2	100	0.87	100	72.3	100	99.5	100	100
D15	53.5	100	0.87	100	25.0	100	96.1	100	99.8
CU	60.5	100	0.86	100	29.3	100	96.7	100	99.8

† Sensitivity
†† Specificity
[1] Positive Predictive Value
[2] Negative Predictive Value

Table 4.8: Table showing efficiency of different colour vision tests including the CAD test in comparison to the Nagel anomaloscope. PPV is the proportion of observers with a diagnosis of deficient colour vision that is correctly diagnosed. NPV is the proportion of observers with a diagnosis of normal colour vision that is correctly diagnosed.

In summary this analysis (Section 4.3.2.1) has revealed a specificity of the CAD test of 100% in comparison to the gold standard test, suggesting that all normals are correctly identified. However 2.5% of colour deficient observers, according to the Nagel test, are incorrectly identified and have normal colour vision according to the CAD test.

The prevalence of colour vision deficiency is of the order of 0.72, slightly lower than that calculated for the other colour vision tests. This is as a result of the inclusion of the random population sample to study the variability within normal colour vision. The relatively high prevalence for colour vision deficiency

observed is due to the fact that most of the subjects examined in this study were referred to us because they had been previously identified as colour defective by failing one or several colour vision tests.

Based on the relatively high prevalence of colour deficiency found in this study, the CAD test reveals 100% PPV (95% CI: 0.97-1) which suggests that all people in our sample diagnosed as being colour deficient are correctly diagnosed. However, 94% (95% CI: 0.85-0.98) of those who passed the CAD test have indeed normal colour vision, i.e. 6% of those diagnosed as having normal colour vision might be incorrectly diagnosed and instead have a colour vision deficiency (according to the Nagel anomaloscope). In the average population, using the values for the screening efficiency in this study, the results suggest that the CAD test provides excellent screening efficiency, and males and females who pass this test are correctly diagnosed with normal colour vision (NPV=99.8% and 100% for males and females, respectively). Those diagnosed as colour deficient are also correctly identified (PPV=100%), unlike the Ishihara test (PPV=41.0% and 3.1% for males and females respectively) which requires further examination with additional tests to confirm the diagnosis.

The rest of the colour vision tests in Table 4.8 were described in Chapter 2 (Section 2.3.2). The overall predictive value (Birch 2001) of screening efficiency (an average of specificity and sensitivity), for all the tests in comparison to the Nagel anomaloscope is: CAD test=98.8%; AO-HRR=97.1%; Ishihara=93.8%; CU=80.3%; D15=76.8%.

Table 4.9 compares the performance, whether pass or fail, on the different colour vision tests. This includes 164 subjects who performed all six tests. It is evident that different colour vision tests often give discordant results (Squire et al. 2005). Similar to the analysis elaborated in Chapter 2, all tests gave the same results for 73 subjects who failed all tests and 17 subjects that passed all tests, so for a total of 55% of the subjects there was full agreement on all tests. Therefore, for 45% of the subjects at least one of the tests gave a contradictory result. As expected, usually the disagreement occurs between the tests with lower sensitivity (the D15 and CU), which pick up moderate to severe colour deficiency, and the tests with

higher sensitivity. This is the case for 64 subjects, who passed one or both of the D15 and CU, but failed the other four tests. The AO-HRR has a higher sensitivity than these two tests, but there are also nine subjects who pass the AO-HRR (and D15 and CU), but fail the Ishihara plates and the Nagel anomaloscope. These subjects are mild deuteranomalous observers who fail the Nagel with RGI values higher than 0.82 but with a midpoint clearly shifted towards higher green values of the red-green mixture scale. On the CAD test, these subjects fail with RG thresholds greater than twice (but less than 3.7) times the average normal CAD observer, and on the Ishihara test at least 1 error is made (that cannot be classed as a misreading) between plates 1-16.

No. of Subjects (total=164)	Colour Vision Test					
	CAD	Nagel	Ishihara	AO-HRR	D15	CU
73	F	F	F	F	F	F
3	F	F	F	F	F	P
16	F	F	F	F	P	F
45	F	F	F	F	P	P
9	F	F	F	P	P	P
1	P	F	F	F	F	P
17	P	P	P	P	P	P

Table 4.9: Comparison of the various colour vision tests, including the CAD test, for the same 164 subjects.

One subject (MED218, see Table 4.2) passed the CAD test and CU tests but failed the other colour vision tests. On the Ishihara test two misreadings were made (for plates 1-16 and 1-25) and on the Nagel anomaloscope this subject failed the 'normal' (having a red-green range of 8 units), and cannot be classified easily into any category but could be described as 'colour weak' on the Nagel. The other colour vision tests failed include the AO-HRR, where one error was made, and the D15, where one crossing and three adjacent transpositions were made.

Additionally, a more thorough analysis reveals two subjects (NC11 and JK16) that pass both the CAD and Nagel tests but fail on the Ishihara plates (1-16 and 1-25 plates). These two subjects also passed AO-HRR and CU tests but did not carry out the D15 test.

One subject (CC125) could be considered to pass the Ishihara plates if misreadings are allowed (only for 1-16 plates). The D15 and CU tests were both passed with no errors. Thus this subject would pass a screening exam based on the Ishihara test that allowed two misreadings. However, this subject failed the CAD (having 2.2 times the normal RG CAD threshold units), Nagel (with a red-green range 16-18) and the AO-HRR tests (making 1 error).

The comparison of Nagel versus the CAD test revealed one protan colour deficient on the Nagel classed as deutan on the CAD (KA264), and vice-versa in another subject (GA35). One further subject satisfied the standard normal observer criteria on the CAD test but might be diagnosed as having deutan colour vision deficiency on the Nagel anomaloscope (DJW144). However, an earlier colour vision report from DJW144 showed that he passed three different lantern tests suggesting no colour vision deficiency.

4.4 Discussion

The standard normal observer on the CAD test, evaluated from 245 subjects, provides an efficient method to determine whether an observer has normal or deficient colour vision. If the individual chromatic thresholds of a subject are within the grey shaded area on the template describing the variability of normal trichromatic vision, the observer is considered to have normal colour vision. The comparison of the new CAD test against the Nagel anomaloscope reveals a very good agreement, with a coefficient of agreement κ=0.97, which means that only a small fraction of the observers tested show divergent results. The agreement in terms of classification of types of colour vision is κ=0.99, i.e., a high value consistent with reliable classification into protan and deutan types.

The results show that the CAD test has a high screening efficiency with respect to established colour vision tests. It has a very high specificity (100%) consistent with the correct detection of minimal colour deficiencies. Plate tests based on the correct identification of a symbol often have significant overlap between responses from normal and colour deficient observers; therefore the diagnosis is usually less certain. As a result, a whole battery of colour vision tests is employed to provide the correct diagnosis and some indication of severity of the colour vision loss. Similarly to the Nagel anomaloscope, the CAD test can grade the severity of a deficiency on a linear scale. The Nagel measures chromatic discrimination ability in relation to the size of the matching range or RGI value. The CAD test scales the extent of colour vision loss by computing chromatic discrimination sensitivity on the CIE-x,y 1931 colour space diagram. The results have shown (see Fig. 4.11) that the measures of RGI and RG thresholds are not equivalent. The properties of the two instruments, particularly the spectral composition of the lights, might explain the differences found between the CAD test and the Nagel. The broadband spectral composition of the display phosphors in the CAD test differs significantly from the narrow spectral lights used in the anomaloscope. The confusion axis tested on the Nagel anomaloscope offers an effective gamut greater than that available on a CRT display (i.e. gamut on monitor cannot be pushed far enough). This observation may explain why severe anomalous trichromats can have RG chromatic thresholds similar to dichromats.

When compared to other colour tests, the CAD test results are in very good agreement with: Ishihara test κ=0.91, AO-HRR κ=0.84, and not that good with: CU κ=0.39 and D15 κ=0.42. Similar findings were obtained in a study presented at ARVO 2004 (Association for the Research in Vision and Ophthalmology). Seshadri et al. (2005) analysed the screening efficiency of the uncalibrated web based CAD test for investigating its validity as a screening test for red/green colour deficiency. She found a high coefficient of agreement of the order of κ=0.93 between the web based CAD test and the standard Nagel anomaloscope results. They found the CAD test to have a specificity of 100% and a sensitivity of 93.33%. The authors also found high agreement with the Ishihara and AO-HRR of κ=0.96 and 0.76, respectively.

A similar display based test is the Cambridge colour vision test (Cambridge Research Systems Ltd. 2002) that employs only spatial luminance noise forcing the observer to rely only on chromatic signals to distinguish the stimulus from the background. Using a version of this test Regan et al. (1994) measured full discrimination ellipses (obtained in 20 minutes) for 48 colour deficient subjects and 41 colour normals. They achieved a good separation between protan and deutan colour deficient subjects, however a limitation encountered (possibly not thoroughly addressed) is the overlap of a small number of anomalous observers (as diagnosed on the Nagel anomaloscope) that have a chromatic discrimination that is within the normal range on their test. In 2003, Ventura et al. (2003) used a later version of this test to produce normative data from a population sample of 75 subjects (age 18-30 years old). The major axis of the ellipse parameters plotted on the u',v' diagram is similar to the results obtained with the CAD test. The advantage of the CAD test, however, is its higher sensitivity and the use of the standard normal template to provide a convenient means of identifying whether an observer has normal or deficient colour vision.

The CAD test can be used to investigate the variability of 'normal' colour vision (by allowing diagnosis of worse and better than the 'normal'). The difference in distribution between male and female populations provides evidence for X-chromosome linked inheritance, and also allows further insight into other possible genotypic factors. It has been suggested that among observers with normal colour vision, the variations are produced by changes in the spectral position of the M and L wavelength cone mechanisms (Neitz and Jacobs 1990). Another possible explanation for the dissimilar distribution amongst males and females is that approximately 15% of the human female population are carriers of X-linked colour deficiencies. Some heterozygotes have been found to exhibit mild abnormalities of colour matching and discrimination. In a previous study, it was found that heterozygotes have significantly larger matching ranges than normal controls (no knowledge of colour deficient relatives) (Jordan & Mollon 1993a). This could account for the two young females that failed the norm on the Nagel anomaloscope in this study. It would therefore be of interest to assess carriers of colour vision deficiencies (with examination of colour deficient sons) to

investigate whether there is an effect on RG chromatic thresholds. Women were shown to have higher RG thresholds than men on the CAD test. If we take 15% of the women tested, this suggests that ~18 women could be carriers of colour vision deficiencies and perhaps account for the upper limit in RG thresholds observed in our sample population. However, there is an alternative possibility that some heterozygous females possess an advantage rather than a disadvantage from the mosaic character of their retina. Potentially, some heterozygote carriers of anomalous trichromacy have the three normal types of cones plus the additional anomalous type that their sons inherit. Such women might be tetrachromatic, enjoying an extra dimension of colour discrimination.

It is important to define accurately the 'normal' criteria on a colour vision test and set pass/fail limits accordingly. As a result, there is a lack of consistency in colour vision testing and a subject applying for a job could be accepted without limitation in one country and rejected in another. For example, this occurs within the JAA member states that use a range of different colour vision tests to examine the colour vision of their applicants (Squire et al. 2005). Additionally, to set appropriate colour vision standards that reflect the pilot's task in the aviation environment, one needs a better understanding of the colour vision requirements in aviation. It is possible, then, to set pass/fail criteria that are safe and relevant to aviation. This applies generally to other occupations, such as the marine and rail transport.

4.5 Conclusions

The combination of temporal and spatial luminance contrast noise has resulted in a test that is rapid to administer and requires no anticipation of the luminance function of individual subjects. The test achieves a good separation of protan and deutan subjects and gives a quantitative measure of colour discrimination. The 16 directions in colour space examined indicate any selective loss of chromatic sensitivity and provide enough information to classify even minimal deficiencies. The template of the standard normal observer on the CAD test provides an

efficient means of identifying even minimal deficiencies that often go undetected in conventional tests. A complete separation between normal colour vision and minimal anomalous trichromats is revealed and supported by the Nagel anomaloscope. The distribution of chromatic thresholds within normal colour vision is in agreement with the basis for gender differences in colour matches (Neitz & Jacobs 1986a). Despite previous concern on introducing temporal luminance noise that it might elevate colour thresholds as a result of increasing noise in the smaller fibres carrying the chromatic signal, temporal dynamic luminance noise has proven a success.

Colour vision, in particular chromatic sensitivity, is a very good indicator of the normal functioning of the retina. Chromatic sensitivity is affected most and earliest in a number of diseases of the retina and the optic nerve. The CAD test can therefore be used to monitor significant changes in both red-green and yellow-blue sensitivity due to progress of disease or the outcome of treatment.

Apart from the different properties of the Nagel anomaloscope and the CAD test that may explain to some extent the disagreements found. It is of interest to attempt to account further for the discrepancy found amongst these two tests. Therefore, a study that investigates the genotype of some observers with particular colour matches on the Nagel anomaloscope is carried out in the following chapter.

5 Colour Vision Assessment in Subjects with Unusual Anomaloscope Matches – a Molecular Genetics Insight into Colour Deficiency

The Nagel anomaloscope is typically used to diagnose the type and severity of colour vision deficiency. The majority of colour deficient observers require either more red or more green light to match the appearance of a spectral yellow field. Some colour deficient subjects accept only very narrow ranges of red/green mixtures that correspond to a red-green discrimination index equivalent to or better than a normal trichromat. More unusual are subjects that are able to match with both an excess of red and an excess of green but do not match all red/green mixture ratios. The purpose of this chapter is to account for this large difference amongst observers in the parameters of the yellow match and to explain the poor correlation between these parameters and estimates of chromatic sensitivity based on colour detection thresholds. A number of unusual colour deficient observers were selected for this study and molecular genetic analysis was carried out to investigate the relationship between predicted and the observed colour vision losses.

5.1 Introduction

Normal trichromats can match any colour by combining three suitably chosen primaries in appropriate proportions. Those that differ from normal trichromats with respect to the proportions of the primaries, may either require that the three primaries be present in unusual quantities, or may only need two primaries. The first type of variation is known as anomalous trichromatism and it is assumed that it arises when three classes of cones are present, one of which contains a photopigment with an anomalous absorption spectrum. The second type of variation is known as dichromatism and this occurs if only two of the three classes of cones are present. The type of deficiency involved causes differences in the chromatic sensitivity that can be identified easily by psychophysical tests of colour vision. Different classes of phenotypes can be identified in this way and these correspond to L-, M- or S- cone deficiency. On the other hand genotype refers to characteristics of an individual's genes. Studies of the empirical relation between a genetic code and an expressed phenotype have lead to improved understanding in the physiological mechanisms underlying many individual differences in colour vision.

Nathans and colleagues (1986a) were the first to use advances in molecular genetics to study human photopigment genes, involving the technique of recombinant DNA (deoxyribonucleic acid). Visual pigments are the light-absorbing molecules that mediate vision with absorption maxima at approximately 420 nm (S-), 530 nm (M-), and 560 nm (L-wavelength sensitive pigment). Structurally, they consist of opsin, transmembrane heptahelical proteins of a single polypeptide chain composed of either 364 (M and L-cone pigment genes) or 348 (S-cone pigment gene) amino acids bound to a chromophore, 11-*cis* retinal (see Fig. 5.1). Photon absorption by the pigment molecules initiates visual excitation by causing an 11-*cis* to all-*trans* isomerisation of the chromophore (see Section 1.3.2.1). The binding socket site of the chromophore in both the cone and rod opsins is located in helix 7 (Fig. 5.1).

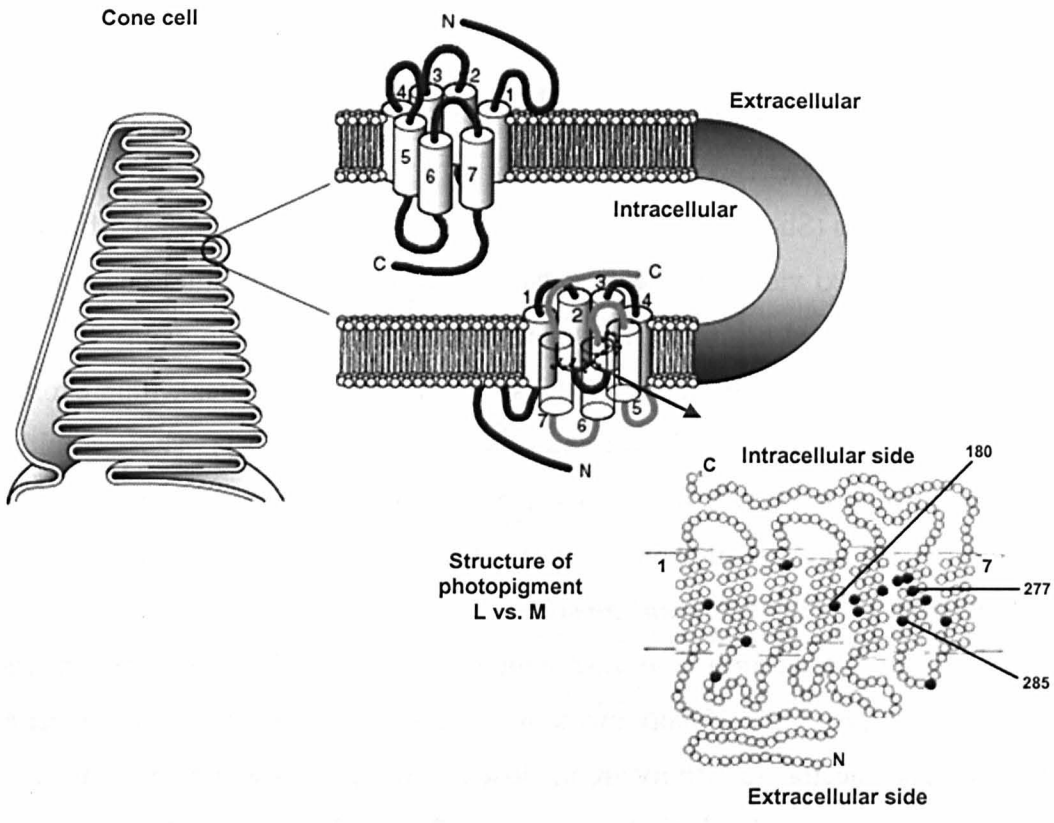


Figure 5.1: Arrangement of visual pigments in a cone photoreceptor. The infolding membrane of the cone outer segment is packed with photopigment molecules. The molecule consists of seven α-helices which span the membrane of the cell surrounding the chromophore, 11-*cis*-retinal. The NH_2 (N) terminus is extracellular, whereas the COOH (C) terminus is intracellular. Shown in the bottom right is the sequence of amino acids that make up this heptahelical molecule. Comparison between the L- and M-cone pigments showing the amino acids that are equal (white circles) and different (dark circles). The substitutions at positions 180, 277 and 285 are believed to contribute most of the spectral difference between the M- and L-cone pigments. Adapted from Sharpe (1999) and Nathans (1986b).

The precise wavelength maxima (λ_{max}) of a given pigment will be determined by the amino acid sequence of the opsin and the interaction of specific amino acids with retinal (Bowmaker 1998). The exact amino acid sequence of opsin 'spectrally tunes' the pigment to a specific wavelength (Asenjo et al. 1994; Neitz et al. 1991).

The genes encoding the S-cone pigments reside alone as a single copy on the q-arm (long arm) of chromosome 7 (Weitz et al. 1992). The genes encoding the L- and M-cone pigments reside on the q-arm of the X-chromosome organised in a tandem array (Nathans et al. 1986b; Vollrath et al. 1988) of up to six exons

(coding regions of DNA) separated by relatively long introns (non-coding regions of DNA) (Fig. 5.3). The S-cone pigment shows only 43 ± 1% analogy in the amino acid sequence in comparison with either the M- or L-cone photopigments. In contrast, the M- and L-cone pigment genes are 96% homologous, differing only in 15 amino acids (Sharpe et al. 1999). The approximate 30 nm spectral difference between these two visual pigments must be attributable to substitutions at these particular amino acid positions. These differences are confined to exons 2-5. The largest shifts in λ_{max} are produced by substitutions at two key sites within exon 5: amino acid positions 277 (~7 nm) and 285 (~14 nm) (Fig. 5.1). These result in spectral shifts of 15-25 nm, the exact value depending on sequences in exons 2-4. Substitutions at the sites in exons 2-4 produce much smaller shifts of less than ~4 nm, and may be responsible for the subtle differences underlying anomalous and normal colour vision. Both in vitro (Asenjo et al. 1994; Nathans et al. 1986b) and in vivo (Sharpe et al. 1998a) methods used to investigate protein sequence variation and spectral sensitivity are in close agreement. These methods have also shown that amino acid substitutions in exon 2 contribute very little to spectral tuning (0.0-0.7 nm) (Asenjo et al. 1994; Sharpe et al. 1998a). More recently, it has been speculated that the amino acid differences in exon 2 are involved in controlling the optical density of the L- and M-cone photopigments (Neitz et al. 1999). The optical density depends on the concentration of the photopigment in the cone outer segment, the length of the cone outer segment and also the extinction coefficient which describes the probability of a photon being absorbed (Wyszecki and Stiles 1982). Changes in the effective optical density of cones cause a broadening of the spectral sensitivity curve away from the absorption peak of the pigment (Fig. 5.2). If two genes that differ only in their exon 2 sequences are expressed, the resulting differences in the optical density of the photopigments may account for some residual colour vision discrimination (He & Shevell 1995; Neitz et al. 1999).

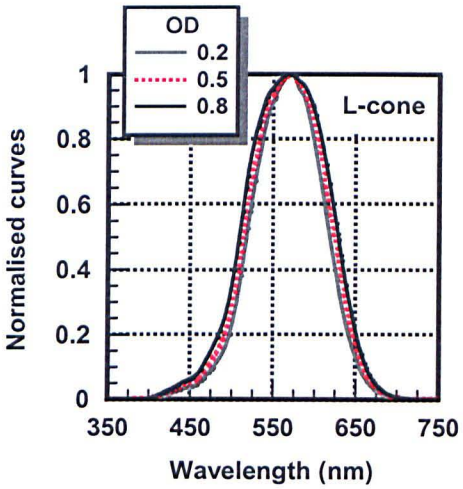


Figure 5.2: Theoretical spectral sensitivity curves for the L-cone sensitive pigment. Changes in the optical density of the cone pigment may occur from amino acid substitutions in exon 2 altering the stability of the molecule or the efficiency with which it absorbs light. For wavelengths near the peak, the optical density qualitatively mimics the difference produced by a spectral shift.

The close similarities between the L- and M-cone sensitive pigment genes suggest evolution from a single gene. Comparison of amino acid sequences suggests that S-cones and rod receptors arose first from a common ancestral receptor. From comparisons with contemporary New World monkeys, which only have two photopigments, it is thought that a long wavelength gene duplicated and diverged to originate the red and green photopigments, at the time when Old World monkeys (trichromatic) separated from New World monkeys (Nathans et al. 1986b; Nathans 1999). The location of the genes for the M- and L-cone pigments on the X-chromosome can account for the larger number of (red/green) colour deficient men compared to women, as men have only one X-chromosome. The head to tail arrangement of the L- and M-cone genes on the X-chromosome is susceptible to mispairing during meiosis, leading to unequal crossing over between gene arrays. If the crossover occurs between genes, this will result in the deletion of a gene from one chromosome and its addition to the other, whereas a crossover within a gene will lead to the production of a hybrid gene that combines regions of the L and M genes into a single gene. Such hybrid genes are thought to

be the genetic basis responsible for the majority of colour vision defects (Nathans et al. 1986b).

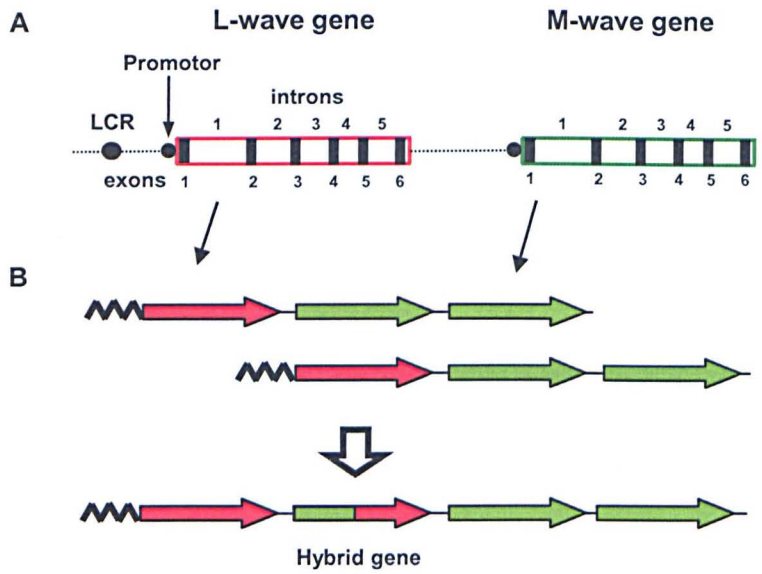


Figure 5.3: Schematic of the tandem array of L- and M-cones on the q-arm of the X-chromosome. (A) The LCR (Locus Control Region) can activate only one of the promoter regions just upstream of a gene (written communication by M. Neitz). The promoters are regulatory units upstream of the transcription site and regulate the rate of DNA transcription into RNA and hence the amount of opsin gene expression (for DNA, upstream is to the left and downstream is to the right). The opsin-coding sequences are divided into exons (black bars). The intron sequences (gaps) are silent or non-coding and usually believed to have no apparent function. (B) Schematic of unequal intragenic crossover that would produce hybrid genes. Adapted from Neitz and Neitz (1995).

Fig. 5.3 shows an example of how the formation of hybrid genes produces pigments with abnormal spectral sensitivity. The spectral sensitivity of the photopigment will depend on which L or M-cone gene the crossover originated from (Neitz et al. 1996). Visual pigment gene arrays for people with normal colour vision have an L gene in the most upstream or leftmost position (Vollrath et al. 1988) and M genes are downstream or to the right of the L gene. Colour normals usually have only one copy of the L-cone gene, multiple copies of M and possibly a number of hybrids (Merbs and Nathans 1992a; Neitz et al. 1993). There is good evidence from studies of gene expression in the retina and from studies involving phenotype-genotype relationships (Hayashi et al. 1999), that only the

first two genes are expressed and have a significant role in colour vision. Whilst colour normals have an L- and an M-cone gene occupying the two most proximal positions of the array from the LCR, anomalous trichromats must have a hybrid gene positioned in either the first or second position. Deutan colour vision arises if a hybrid gene encoding an M-cone gene sensitive pigment occupies second position. Protanomalous observers are missing normal L-sensitive photopigments and colour vision arises from two M-cone genes that differ subtly in spectral absorption properties (Neitz et al. 1991). The M- and L-cone opsin gene can take many forms giving hybrid variants and polymorphisms, and it is this variability underlying the huge differences in red/green colour vision observed within the population. Even amongst people with normal colour vision, differences in colour matching behaviour are associated with individual differences in L-cone pigments (Neitz et al. 1993; Neitz et al. 1995; Winderickx et al. 1992).

By contrast, the S-cone opsin gene sequence is nearly invariant in the human population; however further phenotype-genotype studies are yet to be performed to prove this. Intragenic crossover, the mechanism that permits the frequent manifestation of anomaly in protan and deutan defects, has no analogy in tritan defects. No polymorphisms causing shifts in λ_{max} have been reported so far and only one substitution was found in the coding sequences and exon-intron junctions (Crognale et al. 1999). Three mutations have been established causing amino acid substitutions that perturb the structure or stability of the S-cone photopigment (Weitz et al. 1992), therefore strongly affecting the performance of the S-cone photoreceptors. Further complications lie in the fact that any small variations in λ_{max} would be difficult to dissociate from individual variations in macular pigment and lens density measured psychophysically or *in vivo* (see Chapter 1 and 6).

In the present study it was of interest to determine the genotype of anomalous trichromats and dichromats and correlate this to the variability in colour vision phenotype. The huge variability in Nagel matches, the lack of correlation between midpoint and range of the match and the poor correlation between these

parameters and chromatic discrimination under more normal conditions of illumination were investigated in this study. In particular, we asked what difference in peak spectral sensitivity between the two pigment genes is needed to confer a certain level of red/green colour discrimination.

Clinically, the types of Rayleigh (red-green) colour matches made on a Nagel anomaloscope can characterise the protan and deutan phenotypes. The observer is required to match a spectral yellow primary light to a juxtaposed mixture of red and green primary lights (see Section 2.2.5.1). Normal colour vision observers select an (almost) unique match between red/green mixture ratio and the yellow intensity. Dichromats match the spectral yellow primary to any mixture of red and green primaries by merely adjusting the intensity of the yellow, regardless of the red-to-green ratio. Anomalous trichromats, possessing an anomalous L/M (protanomalous or M-shifted) or M/L (deuteranomalous or L-shifted) pigment instead of the normal (L or M) ones, have matching ranges that are intermediate between those of normals and dichromats, suggesting a less severe form of colour vision deficiency (Pokorny et al. 1979; Sharpe et al. 1999). In addition, a relationship has been found in some deuteranomalous trichromats between the Rayleigh match and the separation of the peak sensitivity wavelengths between a normal L-wavelength gene and an abnormal hybrid gene in second position (Nathans et al. 1986a). It has been suggested that the variability in the Rayleigh matching ranges and midpoints is associated with the nature of the hybrid gene and the anomalous pigment it encodes, with smaller spectral separations (between the anomalous and remaining normal pigment) corresponding to greater severity. The smaller the difference between the absorption spectra between the two adjacent photopigments, the poorer the wavelength discrimination in that region of the spectrum and the more deviant will be the colour matching performance. However, other suggestions have been made to explain cases in which the same phenotype is associated with different genotypes (genetic heterogeneity) and vice versa (He & Shevell 1995; Shevell and He 1997b). Previous studies (Jagla et al. 2003) have shown that a considerable phenotypic variability is associated with any given genotype (e.g. through modifier genes or environmental factors). In 1956, Jameson and Hurvich suggested that the variability among deuteranomalous

subjects' colour matches could be explained by a mechanism beyond the receptoral level (Hurvich and Jameson 1956). A study carried out by He and Shevell (1995) showed that in deutan subjects with possible differences in optical density between visual pigments, as a result of the amino acid sequence in exon 2, effects on colour discrimination are produced. It was shown that a spectral sensitivity difference between the two pigments arising from differences only in optical density (and not in λ_{max}) is sufficient to provide some measure of colour discrimination.

To investigate the relation between chromatic sensitivity, colour matching abilities and genotype revealed from direct DNA sequencing, subjects were selected with extended Nagel anomaloscope results that cross over the normal range, and with narrow ranges either with extra green or red in the match. These subjects also showed mixed phenotypical results in the battery of tests. It is a complex mechanism that makes it difficult to equate genotype and phenotype. The aims of this study can be summarised as follows:

- To correlate genotype predictions with phenotypical behaviour using the Nagel and the CAD tests. To investigate to what extent genetically predicted differences in cone pigment optical density and spectral sensitivity can account for the observed variability in chromatic discrimination.
- To establish the spectral sensitivity of the L-, M-, and hybrid genes or their combination that confers certain levels of red-green colour discrimination.

5.2 Methods and Subjects

From amongst the many subjects examined throughout this investigation (see Chapter 2), 23 subjects that exhibited extreme or unusual anomaloscope matches were selected for the genetic study. Only males have been chosen to take part in this study, as the genetic mechanism in colour deficient females is still less well understood. Further, only subjects with inherited X-linked (red/green) colour vision deficiencies were examined, as S-pigment genes show practically no

variation (Crognale et al. 1999) and yellow/blue deficiency cannot be diagnosed by Nagel anomaloscopy.

Appropriate subjects were identified and diagnosis of the type of colour vision deficiency was established using a whole battery of tests including the pseudoisochromatic Ishihara plates, AO-HRR plates, D15, CU, the Nagel anomaloscope and the CAD test (see Chapter 2 for a description of these tests).

For each subject, the limits of the matching range on the Nagel anomaloscope were determined by averaging two matches. The intensity of the yellow field was recorded at these limits. The mean normal midpoint match is 40 ± 4 and red-green matching range ≤ 7 scale units. Pure green is at 0 and pure red at 73 on the Nagel red/green mixture scale. The number of errors was recorded on the Ishihara and AO-HRR plates. In addition, the classification of the type of colour deficiency was noted with both plate tests and estimated by visual inspection of the D15 result diagrams. The pass/fail criterion of the D15 was one or more isochromatic colour confusions demonstrated across the hue circle in the results diagram (see Fig. 2.3). A MacBeth Easel lamp was used as the illumination source for these tests.

The loss of chromatic sensitivity was re-examined in each of the 23 subjects using the CAD test under conditions that isolate the use of colour signals. Chromatic thresholds were measured for each of the 16 directions along the red-green (RG) and yellow-blue (YB) axes in the CIE colour space. YB and RG thresholds were computed by averaging eight thresholds. For those subjects with higher levels of colour vision loss, chromatic thresholds were measured for directions specifically aligned along the deuteranopic and protanopic confusion axes, respectively for deutan and protan colour deficient observers. This allowed six directions to be averaged along the particular confusion axis therefore increasing measuring accuracy of the level of chromatic sensitivity loss. Units for the YB and RG directions in the CIE colour space are computed in terms of the CAD standard normal observer units. A more detailed description of this is included in Chapters 3 and 4.

Molecular genetic analyses were performed by taking a small blood sample at the health centre of City University from each of the selected 23 subjects. Blood samples were packaged appropriately and transported to the Medical College of Wisconsin in Milwaukee, USA. Genomic DNA was extracted from the blood samples obtained from each subject and the polymerase chain reaction (PCR) was used to amplify segments of L- and M-wave pigment genes separately. For each individual, the relative number of L and M genes on the X-chromosome was thus estimated. PCR products were sequenced to determine variants leading to amino acid substitutions. From the patterns of known amino acid substitutions, the spectral separation between L-wave pigment genes for deutans and M-wave pigment genes for protans were predicted. From studies of spectral ERGs (electroretinograms) on deutan subjects, accurate estimates of the spectral shift produced by amino acid substitutions encoded in exons 2, 3 and 4 of L genes have been obtained (written communication by M. Neitz). Amino acid differences between L- or M-wave pigment genes that might produce consistent differences in optical density were also investigated.

For the full technical description of the molecular genetic analysis carried out refer to (Neitz et al. 1995; Neitz and Neitz 1995).

Table 5.1 shows the results of the 23 subjects selected to examine their genotype. Three groups were previously identified and an ID number was assigned for this study to ease identification.

Subjects 1-9 were diagnosed as deuteranomalous trichromats. Their matching ranges on the Nagel anomaloscope are either larger (≥ 7) or more displaced from the normal midpoint match (40 ± 4) with RGI values ranging from 0.22 to 0.90. Within this group subjects 1-4 are classified as 'extreme' anomalous trichromats (Birch 2003; Walls 1959) because they perform extended matching ranges that crossover the normal matching range but do not match all red/green mixtures. The luminance values on both ends of the Nagel match (L_G for the lower and L_R for the higher limits) are equal and neither high nor low. Subjects 5-9 were diagnosed as 'simple' deuteranomalous trichromats showing large matching ranges displaced towards green on the red/green mixture scale. On the pseudoisochromatic plates

more than 20 errors are made on the Ishihara plates, more than 7 on the AO-HRR plate tests, and both D15 and CU tests are failed. On the CAD test, chromatic sensitivity ranges from 12 to 23 times the 'normal' CAD units along the deutan confusion axis. Both protan and deutan errors were made on the AO-HRR, D15 and City University tests; labelled as 'red-green' in Table 5.1.

The second group, subjects 10-15 were diagnosed as 'simple' deuteranomalous trichromats; all having relatively narrow matching ranges and equal luminance (yellow) values on both ends of the match on the Nagel anomaloscope, i.e. RGI>0.88, they perform fairly well on the pseudoisochromatic plate tests, and pass both the D15 and CU tests; their chromatic thresholds range from 2 to 3 times worse than the SNO for the RG direction.

Finally, the third group of subjects, ID numbers 16-23, were diagnosed as protanomalous trichromats. Their matching ranges on the Nagel all reflect characteristic protan deficiencies matching a low value of luminance (yellow) for the red limit of their range; two subjects matched the full red-green mixture range (subjects 17 and 23). Subjects 18 and 22 are classed as 'extreme' anomalous trichromats (see above). Subjects 20 and 21 have a narrow matching range with RGI values 0.97 and 0.93, respectively and pass both the D15 and CU tests. Protan classification was obtained in almost all these subjects with the Ishihara and AO-HRR test plates, and more mixed protan/deutan results were obtained with the D15 and CU tests. Chromatic thresholds for these protanomalous subjects varied from 6 to 20 times the standard normal CAD threshold along the protan confusion axis. The two subjects, 17 and 23, diagnosed as protanopes according to the Nagel anomaloscope show some residual colour discrimination on the CAD test.

ID	Subject no.	Age	Nagel anomaloscope match			Ishihara		AO-HRR		Farnsworth D15	City University	CAD test		
			Match G - R	Yellow match L_G - L_R	RGI	Class. (no. errors)		Class. (no. errors)				D/P†	RG units	YB units
1	107	27	0 - 49	15 -12	0.33	Deutan	(21)	Deutan	(9)	Deutan	Deutan	D	23.31	1.27
2	235	33	3 - 53	13 -11	0.32	Deutan	(23)	Deutan	(9)	Deutan	Deutan	D	22.21	1.02
3	56	16	8 - 55	13 - 12	0.37	Deutan	(21)	Deutan	(10)	Deutan	Deutan	D	20.52	1.30
4	227	31	0 - 58	14 -14	0.22	Deutan	(22)	Deutan	(12)	Deutan	Deutan	D	16.29	1.14
5	148	22	8 - 30	14 -12	0.71	Deutan	(23)	Deutan	(9)	Red-green	Deutan	D	14.45	0.92
6	151	46	0 - 27	13 -14	0.64	Deutan	(21)	Deutan	(7)	Red-green	Red-green	D	11.76	1.31
7	127	28	0 - 38	13 - 10	0.48	Deutan	(22)	Deutan	(7)	Pass	Deutan	D	11.99	0.98
8	170	20	0 - 7	14 -12	0.90	Deutan	(24)	Red-green	(10)	Red-green	Deutan	D	16.89	0.90
9	225	34	2 - 27	14 - 14	0.66	Deutan	(23)	Red-green	(11)	Red-green	Deutan	D	11.84	1.40
10	249	30	11 - 15	12 - 12	0.95	Red-green	(2)	Red-green	(1)	Pass	Pass	D	2.61	1.00
11	183	40	8 - 17	13 - 12	0.88	Deutan	(13)	Deutan	(3)	Pass	Pass	D	2.97	1.08
12	73	22	14 - 22	12 - 13	0.89	Deutan	(8)	Protan	(0)	Pass	Pass	D	2.41	0.81
13	97	29	15 - 18	11 - 13	0.96	Deutan	(2)	Protan	(0)	Pass	Pass	D	2.71	0.95
14	27	61	15 - 17	14 - 14	0.97	Deutan	(4)	Protan	(0)	Pass	Pass	D	1.83	1.31
15	79	33	15 - 19	14 -12	0.88	Deutan	(3)	Red-green	(1)	Pass	Pass	D	1.92	1.20
16	246	30	55 - 63	7 - 5	0.89	Protan	(23)	Protan	(8)	Protan	Protan	P	19.91	0.92
17	244	31	0 - 73	27 - 0	0.00	Protan	(24)	Protan	(9)	Protan	Protan	P	17.36	0.89
18	75	22	7 - 63	26 - 2	0.24	Protan	(23)	Red-green	(6)	Pass	Deutan	P	14.34	0.97
19	247	24	44 - 64	11 - 4	0.86	Red-green	(22)	Red-green	(4)	Pass	Deutan	P	10.05	0.64
20	240	36	61 - 64	5 - 5	0.97	Protan	(22)	Protan	(4)	Pass	Pass	P	11.33	0.83
21	105	37	60 - 65	5 - 4	0.93	Protan	(15)	Red-green	(4)	Pass	Pass	P	5.90	0.83
22	67	48	31- 56	19 - 7	0.66	Protan	(23)	Protan	(8)	Protan	Protan	P	13.72	1.10
23	248	38	0 - 73	33 - 0	0.00	Protan	(24)	Protan	(9)	Protan	Protan	P	16.75	1.09

Table 5.1: Summary of the clinical test results for 23 subjects selected for further testing.

†D/P denotes chromatic discrimination thresholds measured along the deutan/protan axis.

5.3 Results

The results of the molecular genetic analysis are summarized in Table 5.2. The genetic and the psychophysical data (see Table 5.1) reveal good agreement for the classification of subjects colour deficiencies.

From the genotype results, subjects 1-4, are predicted to be deuteranopic due to having only one L gene and no M gene per X-chromosome (single gene deuteranopes). Therefore the difference in $\Delta\lambda_{max}$ is predicted to be zero. Only the first two genes in the array are expressed in the retina and so colour vision is based on the spectral separation between the pigments encoded by the first two genes in the array (Deeb 2004; Hayashi 1999).

Subjects 5-9 have multiple L and M genes and therefore a difference in $\Delta\lambda_{max}$ based on the first two genes is possible. For deutan subjects with M genes, the M genes are not expressed and the first two genes in the array are both L genes. Therefore differences in these L hybrid genes provide the residual colour vision. The capital letters in Table 5.2 columns 5 to 8 represent amino acids encoded by the pigment genes. The first three letters, TIS or IVY are the exon 2 encoded positions. If one pigment has TIS and the other IVY then the peak spectral separation of the pigments will differ by 2.5 nm.

Subjects 10-13 differ additionally in amino acid sequence in exon 3, the next 5 letters. If the last letter on one gene specifies S and the other A, the two L pigments will differ in λ_{max} by 3.5 nm, resulting in a total separation of 6 nm. In addition, the last three letters are the amino acid polymorphisms in exon 4 of the genes and if they differ, the two L pigments will differ in λ_{max} by 4 nm. For example in subjects 14-15, the amino acid sequences differ in exon 2 (TIS or IVY), exon 3 (S or A), and exon 4 (AIM or TSV) resulting in a predicted spectral separation of the two L pigments of 10 nm. The λ_{max} values for each of the L pigments are given in parentheses after the relevant amino acids for the deutan subjects.

The genotype results show that subjects 16-23 carry only M genes resulting in protan colour deficiency. Subject 16 has only one visual pigment on the X-

chromosome, encoding an M pigment; hence he is a single gene protanope. Subject 17 has one L and two M genes; the two M genes reside in the first two positions in the array thus explaining his protan colour vision deficiency. None of the amino acid differences encoded by exons 2, 3 and 4 of M pigments genes for subjects 17-23 shift the spectral peak of the M pigment, resulting in a predicted λ_{max} separation of 0 nm.

An optical density difference in photopigments is usually associated with differences in amino acid sequence encoding exon 2; this is indicated in Table 5.2. For protan subjects (subjects 17-23) that have M pigments differing at exon 2 encoded amino acid positions (TIS versus IVY), there is a possibility that the pigments differ in optical density. For deutan subjects an optical density difference is possible between pigments but the spectral difference ($\Delta\lambda_{max}$) is probably more relevant and accounts for the specific colour vision abilities. Those subjects, labelled NO or NA, that have only one X-chromosome encoded pigment or for which all of the X-encoded pigments have identical amino acid sequence in exon 2, do not show optical density differences.

ID no.	Subject no.	#L, #M genes	Predicted phenotype	1st L	ds L	1st M	ds M	Predicted ΔΛmax	OD difference
1	107	1L, 0M	deuteranope					0	NO
2	235	1L, 0M	deuteranope					0	NO
3	56	1L, 0M	deuteranope					0	NO
4	227	1L, 0M	deuteranope					0	NO
5	148	3L, 2M	deutan	TISLVAIAIAM (555.5)	IVYMVAIAIAM* (553)			2.5-0	NA
6	151	2L, 1M	deutan	TISLVAISIAM (559)	IVYMVAISIAM (556.5)			2.5	NA
7	127	2L, 2M	deutan	TISLVAIAIAM (555.5)	IVYMVAIAIAM (553)			2.5	NA
8	170	2L, 2M	deutan	TISLVAIAIAM (555.5)	IVYMVAIAIAM (553)			2.5	NA
9	225	2L, 2M	deutan	TVSLVAIAIAM (555.5)	IVYLVAIAIAM (553)			2.5	NA
10	249	2L, 1M	deutan		composite L : t/l i/v s/y m/l v v/a v/l a/s ia m(559-553 or 556.5-555.5)			6-3.5	NA
11	183	2L, 3M	deutan	TISLVAISIAM (559)	IVYMVAIAIAM (553)			6	NA
12	73	2L, 1M	deutan	TISLVAISIAM (559)	IVYMVVVAIAM (553)			6	NA
13	97	2L, 0M	deutan	TISMVAIAIAM (555.5)	IVYMVAIATSV (549)			6.5	NA
14	27	2L, 1M	deutan	TISLVAISIAM (559)	IVYMVAIATSV (549)			10	NA
15	79	2L, 1M	deutan	TISMVAISIAM (559)	IVYMVAIATSV (549)			10	NA
16	246	0L, 1M	protanope			TISMVAIATSV		0	NO
17	244	1L, 2M	protan		IVYMVAIATSV (549)	TISMVAIATSV	IVYMVAIATSV	0	YES
18	75	0L, 2M	protan			TISMVVVATSV	IVYMVVVATSV	0	YES
19	247	0L, 2M	protan			TISMVVIAIAV	IVYMVVVATSV	0	YES
20	240	0L, 3M	protan			TISMVAIATSV	IVYMVAIATSV	0	YES
21	105	0L, 3M	protan			TISMVAIAIAM	IVYMVAIATSV	0	YES
22	67	0L, 3M	protan			IVYLVAIATSV	IVYMVAIATSV	0	YES (unlikely)
23	248	0L, 4M	protan	no sequences for exons 3 & 4, all Ms have IVY in exon 2				?	NA

Table 5.2: Summary of molecular genetic analysis for 23 colour vision deficient observers. **See next page for further details.**

Table 5.2 continued: The single letter is an abbreviation encoding an amino acid which is made up of triplets or codons of DNA sequence. The amino acids found at the polymorphic positions encoded by the first L gene in the array (1st L) and downstream (ds) L genes that encode pigments that are not in the first position in the X-chromosome array are given. For protan subjects, the first M gene in the array (1st M) and ds M genes give the amino acids specified at the polymorphic positions for the M gene at these two positions. It was difficult to sequence the genes for subject 23 (248) and the data available for this subject is therefore incomplete.

Figs. 5.4 and 5.5 show graphically the results obtained on the Nagel anomaloscope and the RG and YB chromatic thresholds respectively for all 23 subjects (Table 5.1). The measurement of colour discrimination, RGI and chromatic thresholds, on the anomaloscope and CAD test respectively, show a relatively poor correlation between these two tests in Fig. 5.6 (see also Chapter 4). On the anomaloscope, subjects 1-4 do not extend the full matching red/green mixture range, however the genotype results reveals deuteranopia with only 1 L pigment present in the X-chromosome. The chromatic thresholds measured on the CAD test for subjects 1-3 extend to the limits imposed by the phosphors of the display. Subject 4 does show some residual colour discrimination on both the Nagel and CAD tests despite the genotype predictions.

Subjects 5-9, have a predicted $\Delta\lambda_{max}$ of 2.5 nm from the genetic results, on the anomaloscope the RGI increases and chromatic discrimination on the CAD test improves. For subjects 10-15, the λ_{max} separation increases to 6 and 10 nm, RGI increases even more and chromatic sensitivity loss is less pronounced. Subjects 10 and 11 have Nagel matching ranges of 3 and 4 respectively, which is in length equivalent to a normal trichromatic match. On the CAD test the chromatic sensitivity loss for the RG axis is approximately twice that of the standard normal observer. Subjects 7 and 8 have the same genotypes at all known functionally relevant polymorphic sites, yet show distinct differences in psychophysical test results on both the Nagel and CAD test (see Fig. 5.7). This indicates that colour vision might be additionally influenced by other currently unknown mutations in the L and M genes or even other genetic or environmental factors.

Subject 16 shows in his genotype only 1 M, and therefore no possible difference in OD, however makes a match on the Nagel anomaloscope of 55-63, (RGI=0.89 or range of 8). In contrast, the CAD results show a severe loss of chromatic sensitivity limited by the phosphor limits of the display. This subject shows by himself the lack of agreement that exists between the Nagel and CAD tests.

Subjects 17-23 show more variable results between the CAD and Nagel tests. However genetics reveal that possibly OD differences between visual pigments may exist and therefore account for this lack of correlation between the two psychophysical tests. Two subjects, 17 and 23, actually show full extended ranges on the Nagel, typical of dichromatism, but their genotype shows 2 or 4 M genes.

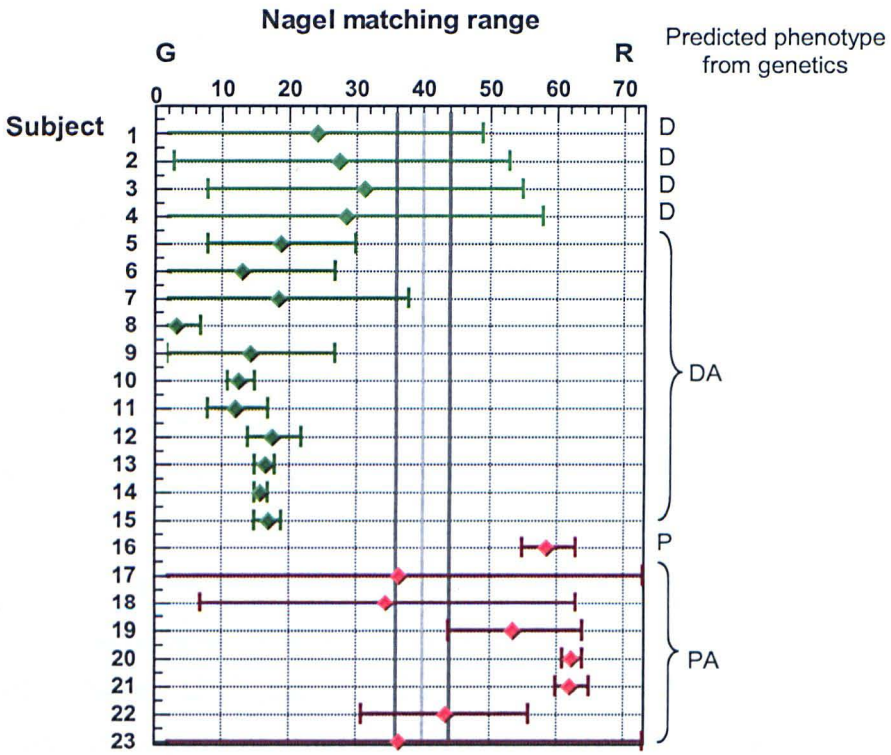


Figure 5.4: Matching ranges and midpoints obtained on the Nagel anomaloscope for 23 subjects that participated in this study. Subjects 1-15 are deuteranomalous trichromats: subjects 1-4 matching ranges cross over the normal matching range (grey lines), and subjects 10-15 have relatively narrow matching ranges. Subjects 16 and 18-22 are protanomalous trichromats, whilst subjects 17 and 23 are protanopic observers according to the Nagel anomaloscope. The predicted phenotype from genetics is shown on the right.

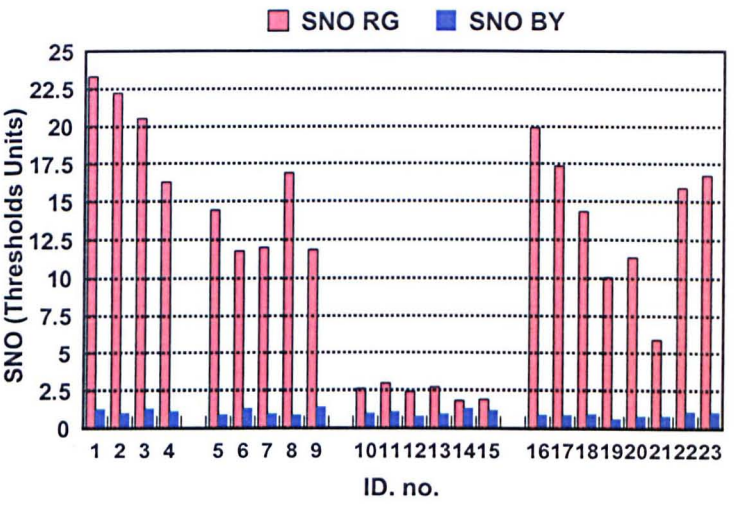


Figure 5.5: Chromatic thresholds measured on the CAD test in 'standard normal observer' (SNO) CAD test units. Thresholds for both RG and YB chromatic directions are shown.

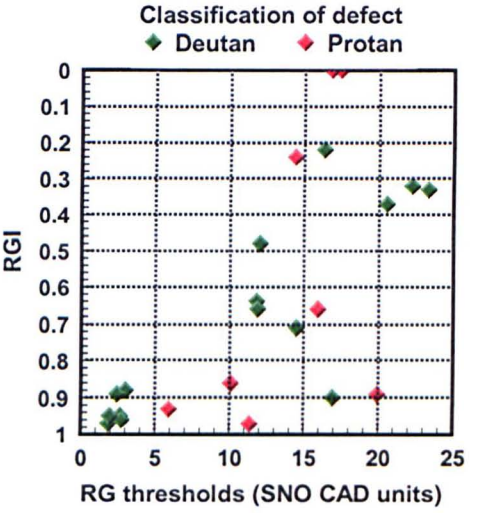


Figure 5.6: Correlation of Nagel anomaloscope (RGI) versus RG thresholds on the CAD test for the 23 subjects examined in this study. A value close to one for RGI reflects a small matching range on the Nagel anomaloscope, i.e. good red-green colour discrimination; while a value close to zero for RGI reflects a large matching range, i.e. poor colour discrimination. A value of RG threshold of 1 reflects the standard normal observer while thresholds above 20-25 reveal severe loss of chromatic sensitivity.

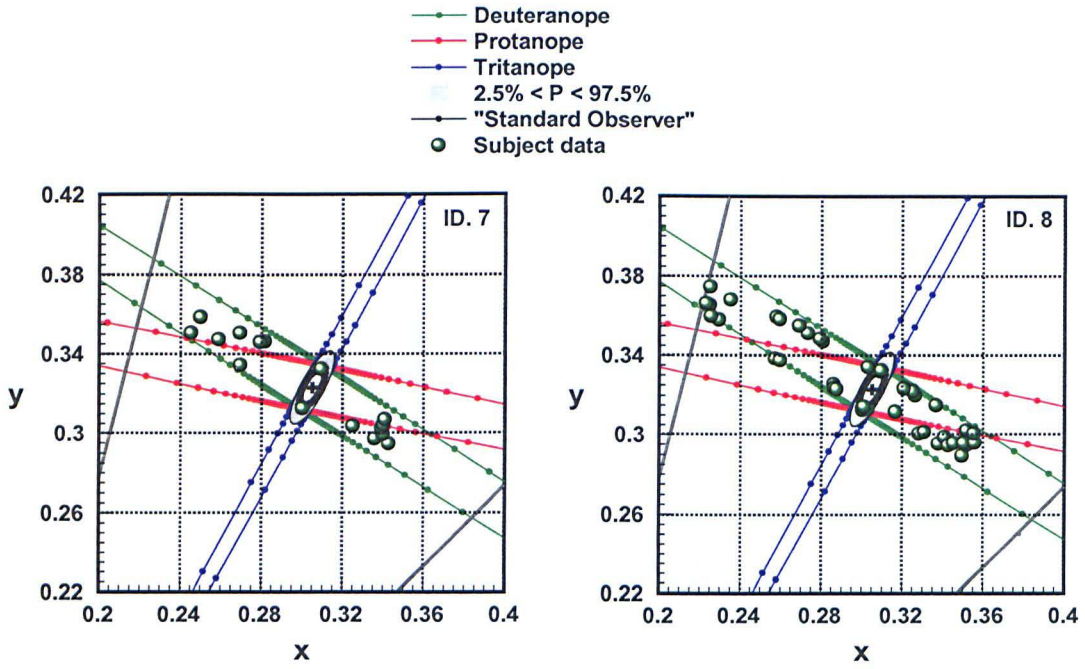


Figure 5.7: CAD data for subjects 7 and 8. These two subjects have the same genotypes with a peak separation λ_{max} of 2.5 nm. On the Nagel anomaloscope: RGI = 0.48 and 0.90, and on the CAD test; RG thresholds = 11.99 and 16.89 for subjects 7 and 8, respectively.

The relationship between the genotype results (Table 5.2) and chromatic thresholds results in units of the SNO on the CAD test can be plotted as shown in Fig. 5.8. (We plot genotype $\Delta\lambda_{max}$ versus chromatic thresholds only, the Nagel range and RGI value gives similar results, however it is thought that the CAD test produces a better correlation due to the discrepancies of subject 16 described above). Here the plot shows the predicted separation in λ_{max} between pigments versus measured RG thresholds. The relationship between these RG thresholds and $\Delta\lambda_{max}$ can be fitted with an exponential function. Subjects with a λ_{max} separation of 0 nm show some degree of colour discrimination. The variability along the ordinate at 0 nm may be due to other factors including optical density changes.

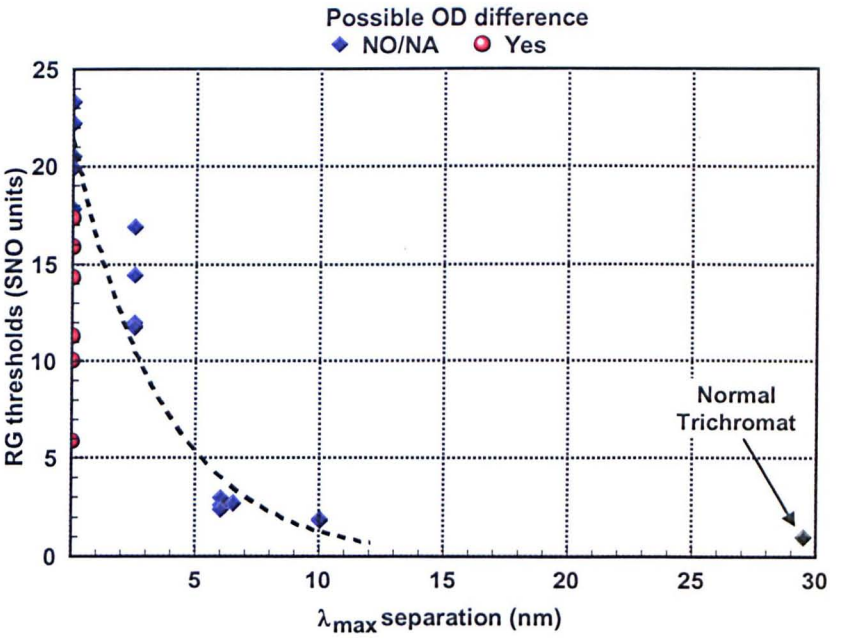


Figure 5.8: Peak separation between visual pigments ($\Delta\lambda_{max}$) versus RG chromatic thresholds. Blue diamonds are subjects 1-15, in which genetic results do not suggest optical density (OD) changes. Symbols in red are subjects with zero peak separation between pigments, but genotype data reveals that differences in OD between pigments is possible and may account for the subjects' colour discrimination.

The analysis of the relationship between RG colour discrimination thresholds and predicted L-, M-cone separation in $\Delta\lambda_{max}$ based on genetics, as shown in Fig. 5.8, suggests that amino acid substitutions that decrease this separation by as much as 8-10 nm may cause no significant loss of chromatic sensitivity and cannot therefore be distinguished from normal trichromats. This is however an extrapolation of the available data and would have to be verified by additional results. This observation does however suggest that amongst the population of normal trichromatic observers, the possession of hybrid genes, encoding visual pigments that cause a reduction in spectral separation between L- and M-cones of less than ~10 nm, may be more common than expected. The supposed separation of 28-30 nm between normal L- and M-cones may not therefore be essential for 'normal' trichromatic vision.

5.4 Discussion

The results described in this chapter show that a reduction in $\Delta\lambda_{max}$ (i.e., the separation in the wavelengths of peak absorbance between the two photopigments) causes a reduction in RG chromatic sensitivity. Variations in absorption spectra are thought to arise from small changes in the protein portion of the photopigment molecule (the opsin) while the chromophore remains unchanged. It has also been suggested that the cone photoreceptors containing the anomalous pigment include a mixture of photopigments molecules from both the M- and L-wavelength receptors. This hypothesis could contribute significantly to the wide variation in colour matching performance and chromatic discrimination found among protanomalous and deuteranomalous trichromats. The results of this study suggest that differences in spectral separation between photopigments are not however sufficient to account for the huge variability in phenotypes found. This is underscored by the observation of individuals with identical genotypes in the known functional polymorphisms who nevertheless show different phenotypes.

Closer examination of the data suggests that other factors beside the investigated genotype may account for additional differences in chromatic sensitivity. The differences in RG thresholds on the CAD test, within each genotype group (i.e. between subjects 5-9 or 10-15) in relation to the intra-individual differences in the test-retest situation (see Chapter 3) suggests that additional parameters must be involved. If the differences in chromatic thresholds within each genotype group are not significantly larger than the typical test-retest differences, one could infer that the outcome is strongly determined by genetic predictions based on $\Delta\lambda_{max}$. If the differences are however significantly larger, one would conclude that other factors beside the investigated genotypes must contribute to the subject's colour vision performance. For example, amongst subjects 5-9, the test-retest differences could be as large as ~2.3 RG standard normal CAD units, such differences fall well outside the expected repeatability error (see Fig. 3.18).

The Nagel performance heterogeneity of anomalous phenotypes is well established. The primaries of the Nagel anomaloscope lie on a common

protan/deutan confusion line, and it was early established that there is little relationship between the mid-point of an anomalous trichromat Rayleigh match and the extent of the matching range accepted (Wright 1946, Hurvich and Jameson 1956; Pokorny et al. 1979). Hurvich and Jameson (1956) suggested that the mean position of the normal and anomalous responses for the Nagel anomaloscope are a function of photopigment peak spectral absorbance whereas the range of responses is due to post-receptoral reductions of the red-green response process. The results of this study suggest that a spectral sensitivity difference among cones arising from differences in optical density (and not in $\Delta\lambda_{max}$) is sufficient to provide the basis for some residual colour vision. In order to assess whether optical changes are possible, in particular amongst the protan colour deficient subjects, Rayleigh matches before and after bleaching might be able to answer whether there is a difference in optical density that provides subjects' colour discrimination (He and Shevell 1994; Shevell & He 1997a).

The results presented here cannot be used to establish whether abnormal post-receptoral gain processes contribute to anomalous colour vision. Pigment λ_{max} separation and optical density together may account for a great variety of anomalous phenotypes, although to fully explain the data obtained one may also need to consider differences in post-receptoral amplification within anomalous trichromats.

5.5 Further investigations

A number of unresolved questions remain and require further experimental studies and modelling work:

- Subjects with genetic arrays of LLM, often show very slight loss of colour discrimination. Despite possible optical density changes being non-applicable (i.e. colour discrimination based only on spectral separation of visual pigments), it would be possible to have additionally quite different optical densities for L and M cones that may not be dictated by the amino acid sequences of the two different pigments in the separate populations of cones. Instead, these variations in optical

density may arise as a variegation that is random (written communication by M. Neitz).

- The difference in optical density hypothesis is important in explaining the residual colour discrimination ability of subject 16 (M gene only subject).
- Further modelling work could be used to determine how $\Delta\lambda_{max}$ between two cone pigments relates to the subject's red-green colour discrimination.
- Genotype data and predictions of $\Delta\lambda_{max}$ may help to constrain the parameters of models of chromatic discrimination.
- Further questions of interest regard the ability to discriminate colours in the visual periphery. It is of interest to know if the genetic specification applies uniformly over the whole retina or if there is any possibility that significant differences in colour discrimination may be found between the fovea and the near periphery. Further studies could be designed to measure colour discrimination thresholds in a number of colour deficient subjects in the near periphery of the visual field. Tritanopes have been reported to exhibit YB discrimination in the periphery in spite of complete absence of S-cone photoreceptors at central vision (Pokorny et al. 1981).

The possible Nagel modelling would be appropriate to understand further the genotype-phenotype relationship.

5.6 Conclusions

The findings of this investigation highlight an important issue in colour vision: how spectral differences between two photopigments actually translate to colour discrimination performance, and how other retinal and cortical factors (such as individual variability in post-receptoral signal gain that precedes the formation of colour-opponent channels) act upon the photopigment differences and influence discrimination. To distinguish anomalous trichromacy and dichromacy it is not enough to look at $\Delta\lambda_{max}$ differences. There are other factors such as optical density and differences in post-receptoral amplification of cone signals that may explain the observed phenotype. When colour vision depends on subtle differences

between the two pigments, as in colour vision defects, relating genotype-phenotype requires consideration of genetic polymorphisms that might affect optical density as well as those that shift λ_{max}. The results obtained in this study suggest that a separation greater than ~20 nm would be sufficient to account for normal colour vision. This implies that the majority of normal trichromats can in principle have hybrid genes encoding their visual pigments providing the RG $\Delta\lambda_{max}$ remains greater than 20 nm.

Predictions based on genetic data in relation to the class and approximate severity of colour vision loss was found to be in agreement with the measured psychophysical data. Changes in $\Delta\lambda_{max}$, photoreceptor optical densities and/or post-receptoral amplification of cone signals appear to be the most important parameters that affect our RG chromatic sensitivity. A complete description of the relative importance of each of these parameters in determining a subject's colour discrimination performance remains a difficult task. The use of information derived from genetic analysis of pigment genes together with psychophysical data on chromatic sensitivity and modelling work could be used to understand and account for the large variability in chromatic discrimination and colour matches observed in both normal and colour deficient observers.

6 Chromatic Sensitivity and Macular Pigment Density in Human Vision

In Chapter 4 it was shown that amongst observers with normal trichromatic vision, there is a greater variability in yellow-blue compared to red-green colour detection thresholds. The aim of this study was to assess whether the variability in yellow-blue discrimination can be attributed, at least in part, to differences in the pre-receptoral filtering of the macular pigment. Although it is well established that selective absorption of blue light by the macular pigment can affect significantly trichromatic colour matches, the extent to which the macular pigment affects colour discrimination sensitivity remains controversial.

In this chapter we have examined how the mean absorption of short-wavelength light by the macular pigment in subjects with normal diets affects chromatic sensitivity. In addition, subjects with increased macular pigment levels, as a result of their participation in a carotenoid supplementation trial, were also examined. Thus we assessed how the variability of yellow-blue and red-green chromatic sensitivity, both foveally and in the paracentral visual field is affected by the peak absorption density and spatial distribution of the macular pigment.

6.1 Introduction

The macular pigment (MP) is composed predominantly of two carotenoids, lutein (L) and zeaxanthin (Z) (Bone et al. 1985), and is of dietary origin (Sommerburg et al. 1998). These carotenoids can be found in various bodily tissues, however their concentration is highest in the photoreceptor axons (Snodderly et al. 1984a; Snodderly et al. 1984b; Sommerburg et al. 1999) in the macula region (Landrum et al. 1999). The MP has band-pass spectral absorption characteristics with peak absorption at ~ 460 nm (see Fig. 1.26) (Bone et al. 1992) thus acting as a pre-receptoral filter that absorbs selectively blue light. The spatial distribution of the MP shows a peak at the fovea and decreases rapidly with eccentricity (Hammond, Jr. et al. 1997a; Hammond, Jr. et al. 1997b). The relative distribution of L and Z vary across the retina with Z dominant in the centre of the fovea and L increasing with eccentricity. The L:Z ratio has been shown to change from 1:2.4 centrally to 2:1 peripherally (Bone et al. 1988). Large, individual macular pigment optical density (MPOD) variations have been reported with extreme differences as large as one log unit (Ruddock 1963). A number of putative roles have been proposed to justify the functional advantages of high MP density levels in the eye (Davies and Morland 2004; Khachik et al. 1997; Kirschfeld 1982; Schalch 1992). These include its antioxidant role that may reduce the risk of developing degenerative diseases of the macula such as age related macular degeneration (ARMD) (Snodderly 1995). The unwanted effects of chromatic aberrations (Reading and Weale 1974) and "blue haze" when viewing distant objects (Wooten and Hammond 2002) can be reduced through selective absorption of blue light. Rod photoreceptors respond well to light in the short wavelength region of the spectrum, but their sluggish responses and extensive spatial summation can cause significant degradation of visual performance. Absorption of blue light by the MP and the corresponding reduction in the effectiveness of rod signals can in principle extend the superior characteristics of cone-mediated vision further into the mesopic range (Kvansakul et al. 2006).

Two aspects of colour vision, colour matching and chromatic threshold sensitivity, have often been investigated in relation to the MP in the eye. Colour

matching involves comparison of two suprathreshold coloured fields, a test field that is matched in perceived colour and brightness with an appropriate mixture of three primary lights of different spectral composition (i.e., the matching field), at least one of which is absorbed by the MP. The measurement of chromatic threshold sensitivity, on the other hand, does not require any judgement of stimulus colour and involves simple threshold detection measurements under stimulus conditions that isolate the use of colour signals (Barbur et al. 1994; MacAdam 1942). Colour matching provides a sensitive measure of differences in the colour appearance of two adjacent fields that reveals variations in the spectral responsivity of cone photopigments as well as differences in their optical densities. These factors contribute to the variability of colour matches, both in dichromats (Alpern and Pugh, Jr. 1977; Alpern 1979) and in normal trichromats (Neitz and Jacobs 1986). In addition, the spectrally selective absorption of light in various structures that precede the photoreceptor pigments in the eye and in particular the absorption of blue light by the MP and the lens also affect colour matches (Van Norren and Vos 1974; Pokorny et al. 1987; Ruddock 1965; Ruddock 1972). The effect MP has on colour matching, particularly when blue/green stimuli are involved, was originally reported by Wright (1929). The measured differences in colour matches made at the fovea and 5° in the periphery, where the density of the MP is reduced significantly, have been used to derive the spectral absorption characteristics as well as the intersubject variation in MPOD (Ruddock 1963; Ruddock 1972).

The discrimination of small colour differences when a colour-defined test stimulus is presented against a uniform background is a measure of chromatic sensitivity and involves the measurement of the smallest difference in the spectral composition of a test stimulus with respect to its immediate surround that is needed to detect its presence (MacAdam 1942). This task is functionally important and is fundamentally different to colour matching. It is generally accepted that colour threshold discrimination measurements involve only two chromatic mechanisms, a red-green (RG) mechanism that compares changes in the outputs of long (L) and middle (M)-wavelength sensitive cones, and a yellow-blue (YB) mechanism that compares changes in the output of short-wavelength

(S) cones against the signals generated in L and M cones. Colour deficient observers have reduced chromatic sensitivity and this reflects selectively the absence or abnormal functioning of one of these chromatic channels. Although existing findings demonstrate beyond doubt that selective absorption of blue light by the MP and the lens affects the colour appearance of the matching field in a colour matching experiment, the extent to which this also affects RG and YB thresholds for detection of colour-defined stimuli is less well understood and remains controversial (Moreland and Dain 1993; Wolffsohn et al. 2000). Improved techniques for the measurement of RG and YB chromatic sensitivity and MPOD make it possible to examine in greater detail the effects of supplementation with carotenoids on MPOD and chromatic sensitivity.

The purpose of this study was to assess how the spatial distribution of the MP changes as a result of supplementation with L and/or Z and to establish whether increased MPOD values reduce chromatic sensitivity in the YB channel. Since supplementation with carotenoids has been shown to improve achromatic vision (Kvansakul et al. 2006) we wanted to establish whether RG chromatic sensitivity may also benefit from increased amounts of MP in the eye.

6.2 Methods

6.2.1 Macular pigment measurements

A visual psychophysics technique based on heterochromatic flicker photometry (HFP) was employed to measure MPOD. Previously, several separate studies have used this method to measure MP (Ciulla et al. 2001; Hammond, Jr. and Caruso-Avery 2000; Hammond, Jr. et al. 2002). This method takes advantage of the fact that the MP is located anterior to the photoreceptors and measures the spectral content incident on the photoreceptors. The optical density measured reflects the sum of the carotenoids present in the fovea. In taking this approach, two conditions are presupposed; first, different concentrations of MPs need to be

present at the foveal and extrafoveal responses and secondly, the response measured is dependent only on the difference in the MP concentration at the two locations (Snodderly 1999). In this estimation, the possible differences in L- and M-cone contributions across the retina and the effects of the photopigment optical density (de Monasterio et al. 1985; Sharpe et al. 1998b) are ignored.

In this study HFP was implemented on a visual display developed in our lab and adapted to provide a rapid and convenient macular assessment profile (MAP) test (Rodriguez-Carmona et al. 2006). The MAP test is based on the use of an optical notch filter to separate the outputs of the three phosphors of the display into two components, one that is absorbed maximally by the MP and is derived only from the blue gun (i.e., the test beam) and the other that is based on a combination of red and green phosphor luminances and consists largely of long wavelength light that is not absorbed by the MP (i.e., the reference beam). The computation of the MPOD reduces to the following:

$$MPOD = \log_{10}\frac{1}{T_{MP}} = -\log_{10}\frac{Rf}{Rp} \qquad (6.1)$$

Where T_{MP} is the transmission of the MP, Rf is the radiance of the blue light needed for the flicker null at the foveal location being measured, and Rp is the radiance for a flicker null at the reference location of 8° in the parafovea, where the effect of the MP is assumed to be negligible.

The luminance of the reference beam is 20 cd/m^2 and its modulation depth is fixed at 20%. The MAP test makes full use of the advantages of visual displays to produce stimuli of varying size at a number of randomised locations, to generate counter-phased sinusoidal modulation of the two stimulus beams (Schalch et al. 2004). The frame rate of the display was 140 Hz and the stimulus modulation frequency was 20 Hz. The high temporal modulation frequency employed ensures that at threshold one isolates the activity of luminance flicker detection mechanisms that rely only on the combined L- and M-cone signals. The stimulus was presented as a short burst of flicker of approximately 0.5 seconds duration and the subject's task was to report the presence or the absence of perceived

flicker. A modified staircase procedure with variable step sizes was then used to measure the mean luminance of the test beam needed to cancel the perception of flicker generated by the reference beam.

The MAP test can be used to measure MPOD along any meridian at a number of specified locations from -8° to $+8^{\circ}$ eccentricity of the visual field. The test stimulus changes from a disc of 0.36° diameter, when presented at the fovea, to a sector annulus when presented at one of five discrete locations on either side of fixation across the horizontal meridian: $\pm 8^{\circ}$, $\pm 6^{\circ}$, $\pm 4^{\circ}$, $\pm 2.5^{\circ}$, $\pm 1.25^{\circ}$, 0° (Fig. 6.1). The width of the test annulus also increases systematically with eccentricity to facilitate the detection and the nulling of luminance flicker. A central spot and radial guides, as shown in Fig. 6.1A, are used to help the subject maintain steady fixation. When the flickering field is placed at an eccentricity less than 1.5° from fixation (see Fig. 6.1A), a static, mirror-image stimulus is generated on the opposite side to facilitate fixation. To avoid Troxler's effect the subject was instructed to blink and/or take breaks from central fixation when attending to the peripheral stimuli (Troxler 1804).

Five, randomly interleaved, repeat measurements were taken at each spatial location investigated and these values are referenced with respect to 8° eccentricity on the assumption that the MP density at 8° is negligible. The test was performed at a viewing distance of 0.7 m and the stimulus was presented only to the right eye. Similar measurements made with the left eye confirmed previous findings (Handelman et al. 1988; Robson et al. 2003), which show good correlation in MPOD values between the two eyes.

The stimulus is always presented at the same location on the screen so as to avoid problems with screen calibration and the spectral transmittance of the notch filter also remains unchanged.

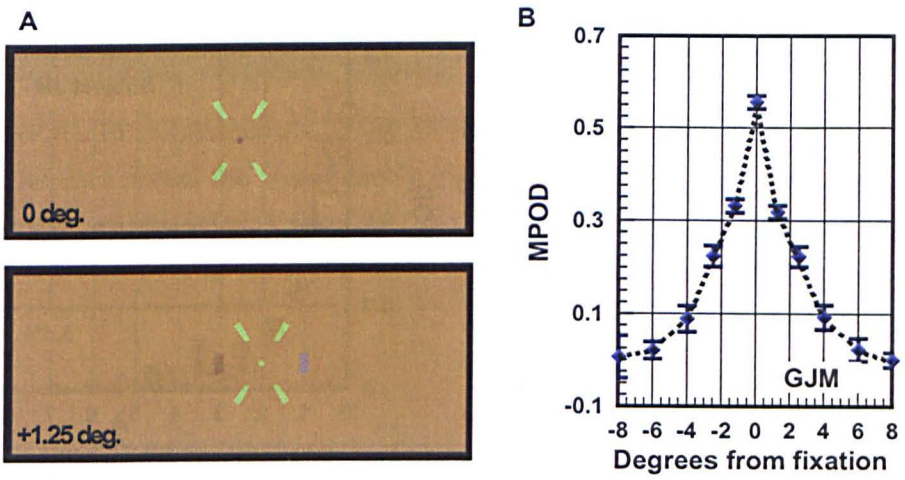


Figure 6.1: (A) Screen dumps showing the appearance of the stimulus employed to measure the optical density of the MP for two locations centred 0° and 1.25° from fixation. At the centre of the fovea, the stimulus was a small disc and subtended 0.36°. For 1.25° eccentricity the stimulus was a sector annulus and was presented simultaneously on both sides of the fovea, an arrangement that facilitates steady fixation. (B) Example of MPOD profile obtained by measuring MP optical density at a number of retinal locations up to ±8°.

The MAP test was compared recently against MP measurements using the well-established Moreland anomaloscope technique modified for motion photometry (Moreland and Kerr 1979; Moreland 2004). Since the latter instrument employs a narrow, short-wavelength beam with a peak output at 460 nm, the technique measures the peak optical density of the MP. The MAP test, on the other hand, employs a broader, short-wavelength beam and this may slightly underestimate the peak MPOD (Moreland et al. 2001). Following correction for peak optical density, the results obtained on the two instruments were found to be in close agreement as shown in Fig. 6.2. In order to assess the agreement between the two measurements for each subject, a line was fitted to the logarithm of both sets of values. Both gradients agree fairly well. The relative or percentage error between these gradients, for each subject, is also shown in Fig. 6.2.

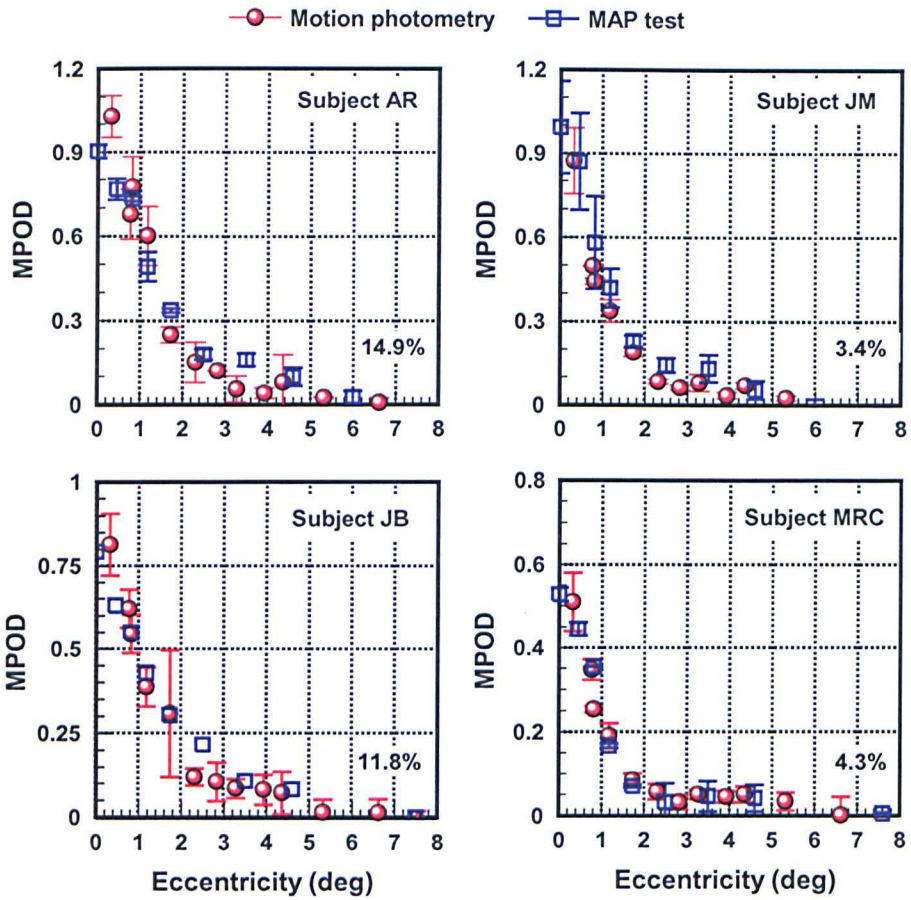


Figure 6.2: A comparison of spatial profiles of macular pigment density measured using motion photometry and the MAP test in four subjects. Each estimate is an average of two measurements carried out on the two methods; error bars indicate ±SD. The two independent estimates of MP optical density profile are in good agreement in each of the four subjects. The percentage error in the gradient (for a log scale) is given.

6.2.2 Measurement of chromatic sensitivity

Chromatic detection thresholds were measured using the CAD test for a number of discrete directions in colour space in order to describe the sensitivity of both the RG and the YB chromatic channels (see Chapter 3 for further details).

Fig. 6.3A shows actual data points for a typical normal trichromat measured at the fovea (coloured symbols) and 3° in the periphery (grey squares). The mean RG and YB thresholds are computed by averaging the chromatic signal strengths

measured in each of the eight, closely packed, hue directions that describe the orientation of each chromatic channel. Examples of RG and YB thresholds measured at the fovea and 3° in the periphery for one subject are shown in Fig. 6.3B. The data reveal the significantly higher thresholds for the YB channel, both at the fovea and 3° in the periphery.

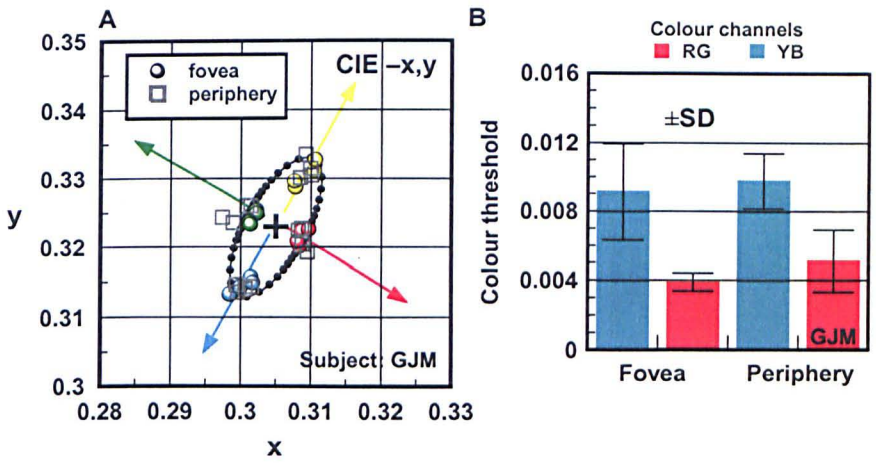


Figure 6.3: (A) Colour detection thresholds for both yellow-blue (YB) and red-green (RG) discrimination measured with the CAD test. Typical results for one normal trichromat. The cross at the centre plots the (x, y)-chromaticity of the uniform background field and the coloured symbols show the thresholds measured in the yellow, green, blue and red directions. Measurements were taken at the fovea (solid discs) and 3° in the periphery (outline squares). The dotted ellipse plots the mean thresholds for the 'standard' CAD observer computed by averaging results for 237 subjects with normal colour vision (see Chapter 4). (B) Bar plot showing typical results for one subject at the fovea and in the periphery. The error bars show ±SD.

6.2.3 Subjects

The subjects examined in this study were recruited initially for a larger carotenoid supplementation trial (Köpcke et al. 2005; Schalch et al. 2005) which began in 2002 and involved a total of 92 participants. Our participation in this study only started in *phase II* (see Table 6.1). The participants received supplementation consisting of L, Z, a combination of the two carotenoids or placebo for twelve months. The subjects were male, Caucasian and aged in the range 22-39 years.

The tenets of the Declaration of Helsinki were observed, and the study had the approval of the research and ethical committee of City University. All subjects gave informed consent before and during participation. A thorough optometric examination was carried out on each subject and no clinically detectable signs of ocular disease or other abnormalities were found.

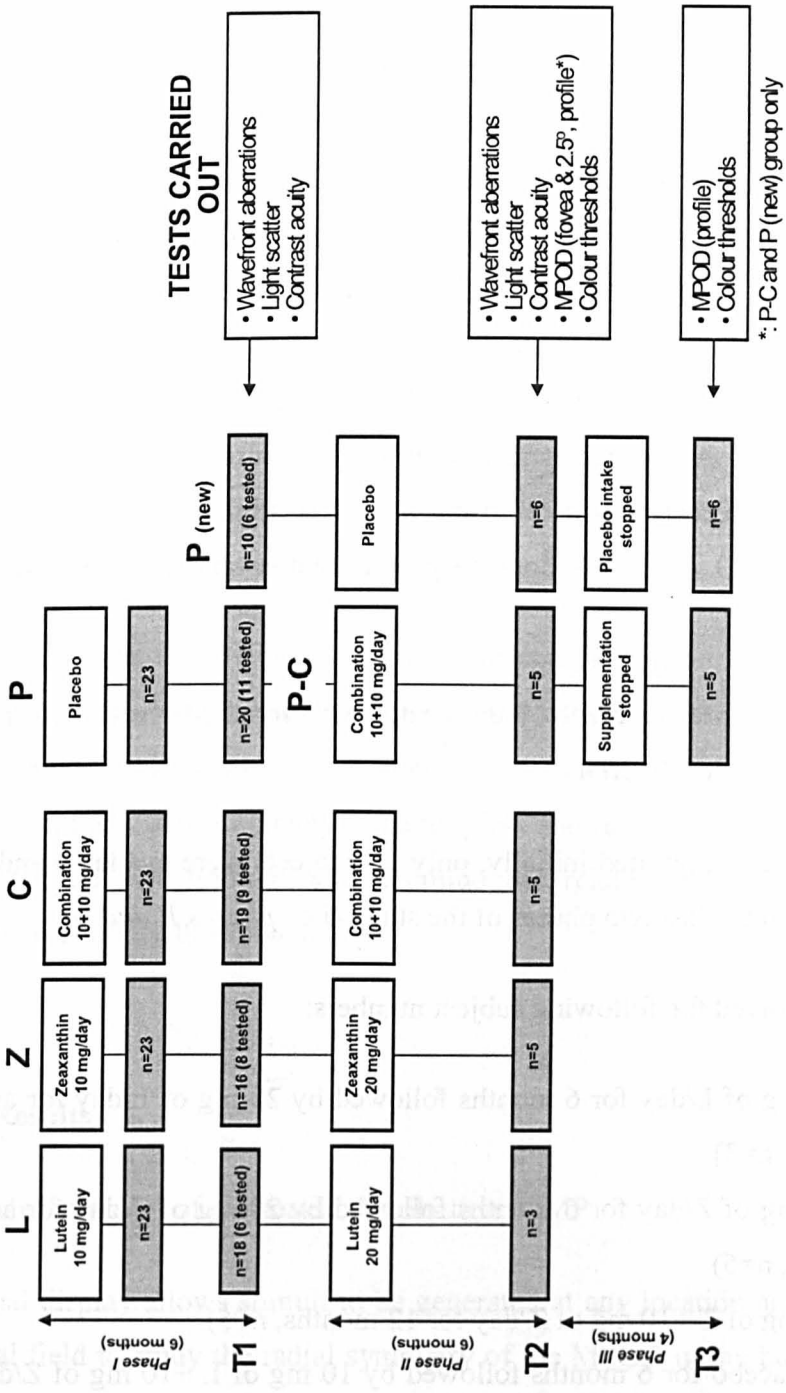


Table 6.1: Schematic diagram showing the various subject groups, the time course of supplementation with carotenoids and the various tests of visual performance carried out at key stages during the trial. Measurements of rms wavefront aberrations, contrast acuity, scattered light and MP density at the fovea and 2.5° in the periphery were carried out at the end of phase I (time T1), at the beginning of phase II (for the new placebo group) and at the end of phase II (time T2), for all subject groups. Colour thresholds and MP density profiles from –8° to +8° were measured at the end of phase II and phase III for the P-C and P groups only (time T2 and T3, respectively). The various groups are as indicated in the table together with the amount of daily intake of carotenoids during supplementation. At the end of Phase I, the P-C group was formed from enrolling all available subjects from the initial placebo group. The P-C group was then given combined L and Z supplementation for six months and 10 further subjects were recruited to form the new placebo group.

A number of optometric and ophthalmologic tests were performed at the end of *phase I* (time T1) and *phase II* (time T2) on each subject group, as shown in Table 6.1. These tests included the measurement of wavefront aberrations, contrast acuity, scattered light and MPOD. The additional measurements needed for this study were made at the end of *phase II* (time T2) and four months after supplementation stopped, at the end of *phase III* (time T3), and involved measurement of MPOD profiles and YB and RG chromatic discrimination thresholds. At time T3 chromatic thresholds were measured at the fovea and at an eccentricity of 5°. At the same time additional measurements were made at the fovea and at an eccentricity of 3° using a smaller stimulus, as described in the caption to Fig. 6.3. This was done to provide better spatial localization of the stimulus.

The MAP test for assessment of the spatial profile of MP was only available during the *phase II* of the trial.

Of the 92 subjects recruited initially, only 23 subjects were available and willing to participate in the last two phases of the study (i.e., *phases II and III*).

The study involved the following subject numbers:

- L (10 mg of L/day for 6 months followed by 20 mg of L/day for another 6 months, n=3)
- Z (10 mg of Z/day for 6 months followed by 20 mg of Z/day for another 6 months, n=5)
- C (10 mg of L +10 mg of Z/day for 12 months, n=5)
- P-C (placebo for 6 months followed by 10 mg of L +10 mg of Z/day for 6 months, n=5)
- P (placebo for 10 months, n=6)

For the second six months, during *phase II* of the study, the dose was doubled for the L and Z groups. The subjects in the P group at the end of *phase I* became part of a new group, the P-C group, and received combined supplementation for six

months. More subjects were recruited and formed the new placebo group. This made it possible to compare the P-C group after only six months of supplementation with the new P group.

This arrangement enabled us to make group comparisons of changes in MP density and chromatic thresholds as a result of supplementation at the end of *phase II* (T2), and four months after the supplementation was stopped (at the end of *phase III*, time T3).

6.2.4 Data analysis and statistics

Analysis of variance (ANOVA) was used to compare mean estimates between groups of subjects. All analyses were performed with S-PLUS® 6.2 for Windows (Professional Edition; Insightful Corporation, Seattle, USA) by W. Köpcke (University of Münster, Germany). Scatter plots and regression analysis was used to confirm the key findings and examine the relation between MPOD and chromatic sensitivity thresholds.

6.3 Results

6.3.1 Measurement of the spatial distribution of MP

The visual display allows stimuli to be generated at any location and meridian in the visual field to study the radial symmetry of the MPOD in the human eye. Fig 6.4 shows the overall spatial distribution of the MP in the retina measured in two subjects. Data was collected every 30° meridian up to 360° at the following eccentricities: ±8°, ±6°, ±4°, ±2.5°, ±1.25°, 0°. The profiles reveal radial symmetry in agreement with other techniques, such as autofluorescence (Robson et al. 2003) that have been employed to measure the spatial distribution.

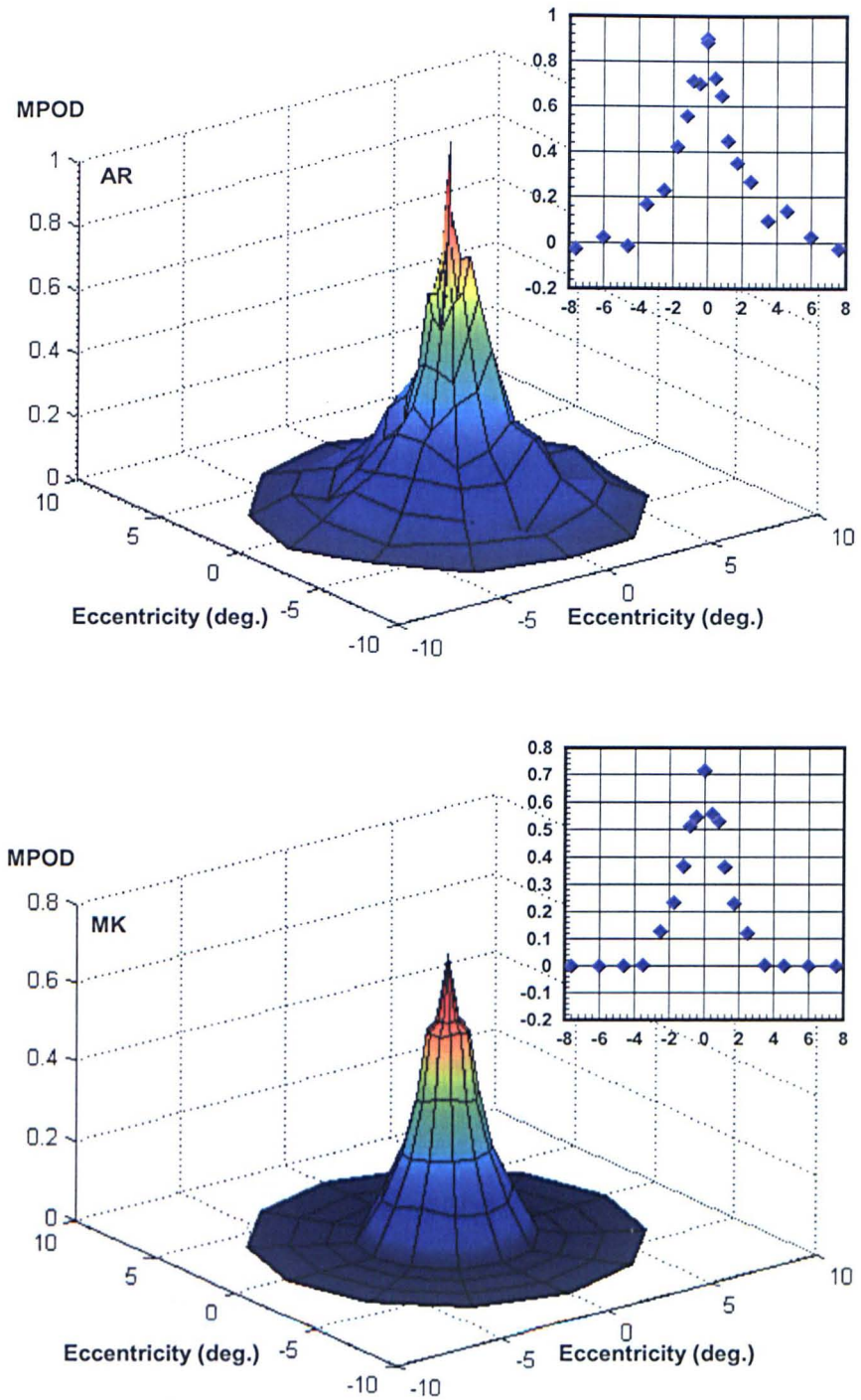


Figure 6.4: 2-D spatial profiles of the MPOD distribution measured using the MAP test for two normal colour vision subjects plotted using MATLAB (The Mathworks, Inc., Massachusetts). Interpolation of the vertex colour values determines the colour of each surface face. The insert shows data measured along one meridian only. Note the 'shoulder' in the profile; a similar observation was made previously (Berendschot & Van Norren 2006). I would like to acknowledgment Anthony Robson and Marale Kianpour for their participation and patience in these experiments.

6.3.2 MP group comparison

The first set of measurements on subjects were carried out after twelve months of supplementation (at time T2), with doubled supplementation for the second six months of phase II (see Table 6.1).

Fig. 6.5 shows MPODs measured in all subject groups, at the end of phase II of the study. For each group that received supplementation, the results show a significant increase in MP density ($p<0.05$ for the C and P-C groups and $p<0.07$ for the L and Z groups; see Appendix E) by comparison with the placebo group. There was no statistically significant difference between the P-C group (after six months of supplementation) and any of the other groups (that received supplementation for 12 months). This result confirms previous findings showing that a diet rich in L and Z can cause increased levels of MP, at least in some individuals (Berendschot et al. 2000; Hammond, Jr. et al. 1997a; Landrum et al. 1997), and that increased MPOD levels could still be detected for at least six months after the diet was stopped (Hammond, Jr. et al. 1997a). The data shown in Fig. 6.5 were measured both at the fovea (A) and 2.5° in the periphery (B) using the new MAP test (see Fig. 6.1). Section C plots the actual increments in MPOD for each of the supplemented groups with respect to the placebo group, both for the fovea and for the periphery.

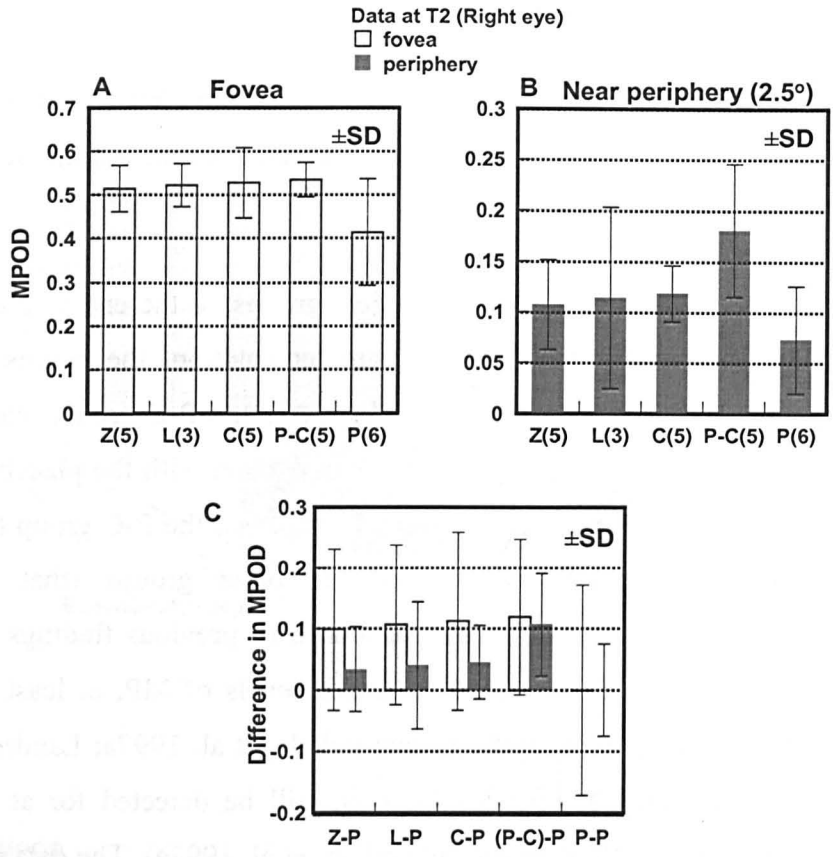


Figure 6.5: Estimates of mean MPOD for each group of subjects at the end of phase II of the supplementation study (see time T2 in Table 1). Measurements were made at (A) the fovea and (B) 2.5° in the periphery. The error bars indicate inter-subject variability in each group. (C) Shows a plot of the actual differences in mean MPOD values for each supplemented group with respect to the placebo group.

In order to evaluate the effect supplementation might have on the spatial distribution of the MP, measurements were made at a number of discrete eccentricities in the range +8° to -8°. These additional measurements were carried out only in the P-C and the P subject groups at time T2 and at time T3, after four months without supplementation. Fig. 6.6C reveals the clear differences in the spatial distribution of MP (measured at T2) for the supplemented and the placebo groups. The increase in MPOD values is significant ($p<0.025$) in the P-C group when compared to placebo (see Appendix E). The measured MPODs were converted to percentage transmittance values in order to express the differences between the P-C and the P groups as a percentage change in transmitted light. These comparisons computed for time T2 (after six months of supplementation)

and at time T3 (four months after supplementation was stopped) are shown in Fig. 6.6D. The results show clearly that the percentage reduction in transmitted blue light as a result of supplementation extends almost uniformly in the paracentral region around the fovea over a visual angle of ±4°. The percentage reduction in the transmission of blue light is significantly less at time T3 after four months with no supplementation. Fig. 6.6A,B show the MPOD profiles for the P-C and the P groups, respectively at both T2 and T3. Comparison of the spatial MPOD profiles for the P-C group reveals a significant decrease in MP density after four months with no supplementation. In contrast, the same measurements show no significant changes in the P group.

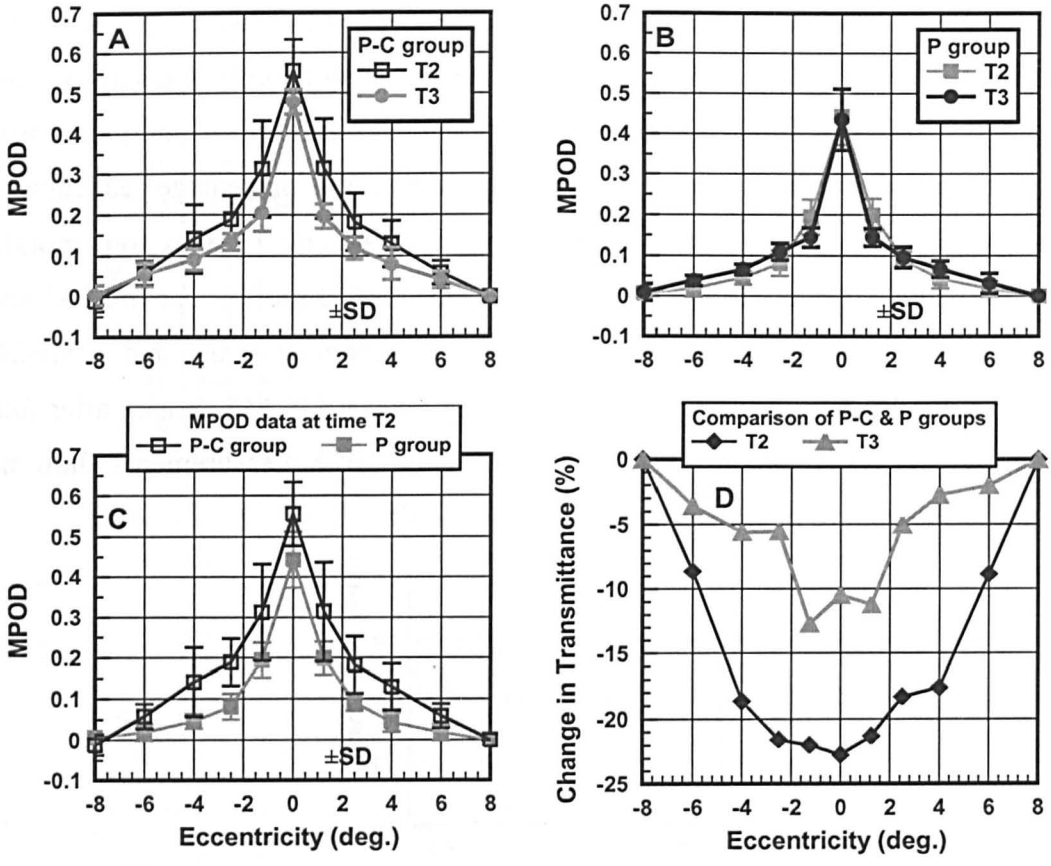


Figure 6.6: Comparison of MPOD spatial profiles for the P-C and the P groups measured along the horizontal meridian from +8° to –8° at different times during the study. (A) Shows data measured in the P-C group at the end of phase II (time T2) and after 4 months with no supplementation (time T3). (B) Similar data for the P group. The error bars indicate the inter-subject variability (±SD). (C) Shows MP profiles for the P-C and P groups measured at the end of phase II of the study (time T2). (D) Shows the percentage change in the transmittance of blue light when the MPOD values in the P-C group are compared with the corresponding measurements in the P group. The diamonds show the difference between the P-C and P groups at time T2 and the triangles show the percentage difference between the same groups, but at time T3.

6.3.3 Chromatic sensitivity versus MPOD

In order to establish the possible effects of MP density on chromatic discrimination sensitivity we evaluated the correlation between the measured RG and YB colour thresholds and the peak density of the MP for all subject groups. The data examined were measured at time T2 and are shown in Fig. 6.7. YB thresholds are plotted in the top two sections labelled A & B. Sections C and D

show similar data describing RG thresholds. A and C show results measured at the fovea and the corresponding data measured 2.5° in the periphery are plotted in B, D. The 15x15 check array employed to measure colour detection thresholds (see Chapter 3) subtended a visual angle of 2.82° at the fovea and 4.23° in the periphery. The coloured stimulus consisted of 5x5 checks and subtended a visual angle of 0.94° and 1.41°, respectively. The larger stimulus size selected for the measurement of colour thresholds at 5° eccentricity provides some compensation for the known reduction in chromatic sensitivity with eccentricity (Abramov et al. 1991; Boynton et al. 1964). The results show little or no correlations between the level of MP in the eye and the corresponding RG or YB colour discrimination thresholds.

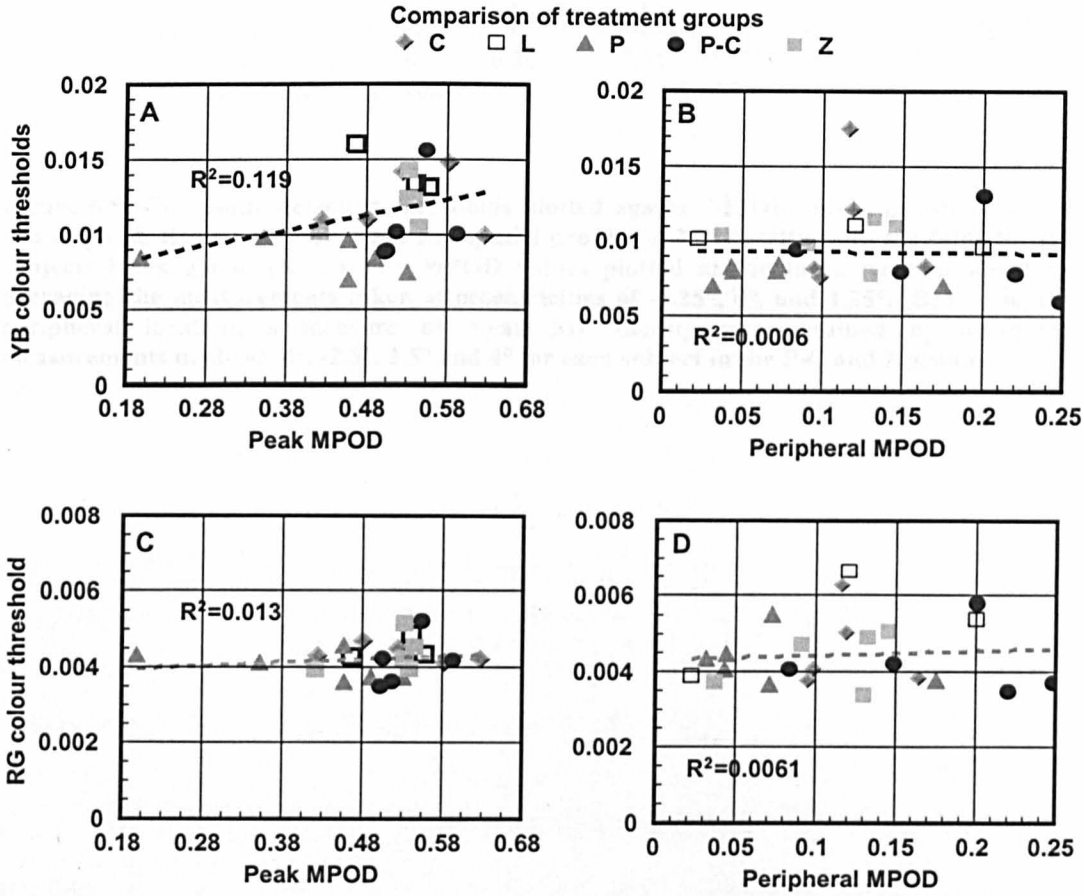


Figure 6.7: (A, C) Chromatic detection thresholds measured at the fovea and (B, D) 5° in the periphery plotted against the available MPOD values at the end of phase II. YB thresholds are shown in sections A and B and RG thresholds are shown in sections C and D.

Similar results were found four months after supplementation was stopped (at time T3) when additional measurements were made in the P-C and P groups. The additional colour discrimination thresholds measured at time T3 employed a smaller stimulus size, of 2°, which moved through a visual angle of ±1°, centred both at the fovea and at an eccentricity of 3°. The spatial localization of the coloured stimulus was improved in order to reveal the possible effects on YB colour thresholds caused by the greater absorption of blue light by the MP at the fovea. Full spatial profiles of MP density were also available for each subject at time T3. This made it possible to obtain mean MP estimates by averaging measurements taken at 0° and ±1.25°, for the fovea, and at ±2.5° and ±4°, for the periphery. The new RG and YB colour thresholds measured under rigorous conditions and with smaller coloured stimuli that provide better spatial localization are shown in Fig. 6.8. The results illustrate convincingly the absence of any significant correlation between the level of MP in the eye and the corresponding RG or YB colour discrimination thresholds.

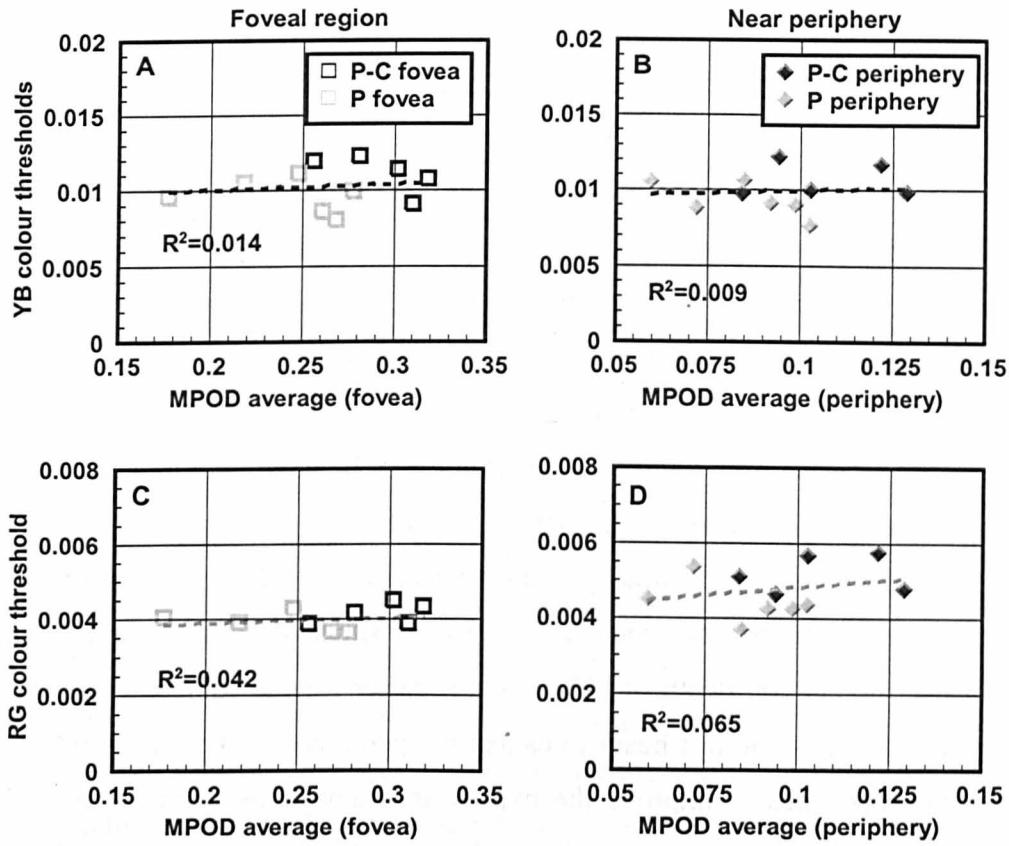


Figure 6.8: Chromatic detection thresholds plotted against MPOD values measured at the end of phase III of the study when full spatial profiles of MP density were available for the subjects investigated. (A, C) The MPOD values plotted at the fovea were obtained by averaging the measurements taken at eccentricities of -1.25°, 0°, and 1.25°. (B, D) For the peripheral location, a measure of mean MP density was obtained by averaging measurements made at -4°, -2.5°, 2.5° and 4° for each subject in the P-C and P groups.

6.4 Discussion

This investigation exploits the advantages of two new techniques, one developed to measure the spatial distribution of the MP and the other to assess RG and YB chromatic sensitivity with stimulus conditions that isolate the use of colour signals (Barbur 2004; Rodriguez-Carmona et al. 2005). The MAP test measures MP density with respect to an eccentricity of 8°, a value considerably larger than other flicker cancellation tests that reference MP changes with respect to 5° or 6° eccentricity (Hammond, Jr. et al. 2002; Wooten et al. 1999). Although the differences between the 5° and 8° eccentricity may be negligible in subjects with normal levels of MP density, the larger, 8° reference is important for use in carotenoid supplementation studies when the accumulation of MP extends to more distant regions around the macula. The use of sinusoidal flicker and a small luminance modulation depth of 20% make it easier for the subject to adjust the mean luminance of the test beam to cancel the perceived flicker in the reference beam and this tends to improve the overall accuracy of the measurement. The findings from this study confirm other earlier reports suggesting increased levels of MP in the eye as a result of supplementation with carotenoids (Sommerburg et al. 1998). In addition this study reveals the extent to which MP accumulates in more distant areas around the macula causing an almost uniform reduction in the amount of blue light reaching the retina within the centre ±4°. Previous studies suggest that the sparing of the very central area of the macula in a variety of macular degenerations is often attributed to the presence of high levels of MP in the spared region (Weiter et al. 1988). The mechanism that contributes to this outcome is probably the absorption of higher energy, short wavelength photons and the more effective quenching of reactive oxygen intermediates that arise from the photo transduction process. The significant increase in MP density at and around the macula region following supplementation with L and Z implies that a larger region of the retina may receive greater protection against ARMD. The implicit claim of this hypothesis rests on obtaining further evidence to validate previous reports suggesting that the size of the spared retina in annular maculopathies relates to the amount and spatial extent of the MP in the eye (Weiter et al. 1988).

The principal aim of this study was to investigate the extent to which the increased variability in YB colour detection thresholds in normal trichromats reflects the large differences in MP density reported in the normal population (Ruddock 1963). The CAD test offers unique advantages to quantify both YB and RG chromatic sensitivity both at the fovea and in the periphery where MP density is significantly reduced. The results show that although both MPOD levels and YB colour thresholds show large intersubject variability, no significant correlation exists between these two variables suggesting that the increased variability should be attributed to other independent mechanisms. This finding was surprising given the well-established effects absorption of blue light by the MP can have on trichromatic colour matches. In order to explain these findings we have examined how cone photoreceptor contrasts change at threshold and how the MP affects these curves. Fig. 6.9A shows the cone contrasts computed for the mean threshold discrimination ellipse (the black line shown in Fig. 6.3A). The cone contrasts are calculated by evaluating the signal generated by the uniform background and the corresponding test stimulus, as a function of its hue direction in the CIE-x,y chromaticity diagram. This computation employs the known spectral responsivity functions of cone photoreceptors (Smith and Pokorny 1972) and knowledge of the spectral radiance data of the test and background fields. No other assumptions were involved.

For example, the M cone contrast signal is given by:

$$C_M = \frac{M_{test} - M_{bkg}}{M_{bkg}} \qquad (6.2)$$

where test and bkg stand for test and background field, respectively. M_{test} and $M_{bkg}$is given by: $\sum_{380nm}^{780nm} R_{test}(\lambda)g(\lambda)d\lambda$ and $\sum_{380nm}^{780nm} R_{bkg}(\lambda)g(\lambda)d\lambda$ where R_{test} and R_{bkg} are the spectral radiance distribution of the test and background fields, and $g(\lambda)$ is the M-cone spectral sensitivity function.

The dotted lines show the expected cone contrast changes when the spectral composition of the light from the display is modified by a filter corresponding to

the MP spectral absorption template for a peak optical density of one log unit, a value significantly larger than those measured experimentally in this study. This computation is very informative in that it demonstrates clearly the independent functioning of the two chromatic channels at threshold (Guth et al. 1980; MacLeod and Boynton 1979), with the RG channel mediating colour discrimination over most of the range. The RG colour channel is formed by differencing of L- and M-cone signals and the YB channel is driven mostly by the S-cone signal which is subtracted from the sum of M- and L-cone signals. The simple model of colour vision predicts little or no change in YB colour discrimination when a MP filter with a peak density of one log unit is placed in front of the eye, as shown in Fig. 6.9B. Interestingly the RG colour signal is actually increased for some hue directions, mostly as a result of reducing selectively the uniform background signals as 'seen' by M- and L-cone photoreceptors without having an equivalent effect on the longer wavelength light of the test stimulus for RG discrimination. This improvement in RG chromatic sensitivity is probably small and the hue directions that are affected most were not investigated experimentally. Although the experimental findings and the predictions of the threshold colour discrimination model show convincingly that YB colour thresholds are not affected even by large amounts of MP in the eye, one could reasonably expect that at relatively low light levels when the signal to noise ratio in S-cones becomes small, YB colour discrimination thresholds may no longer be independent of MP in the eye (see Fig. 3.10).

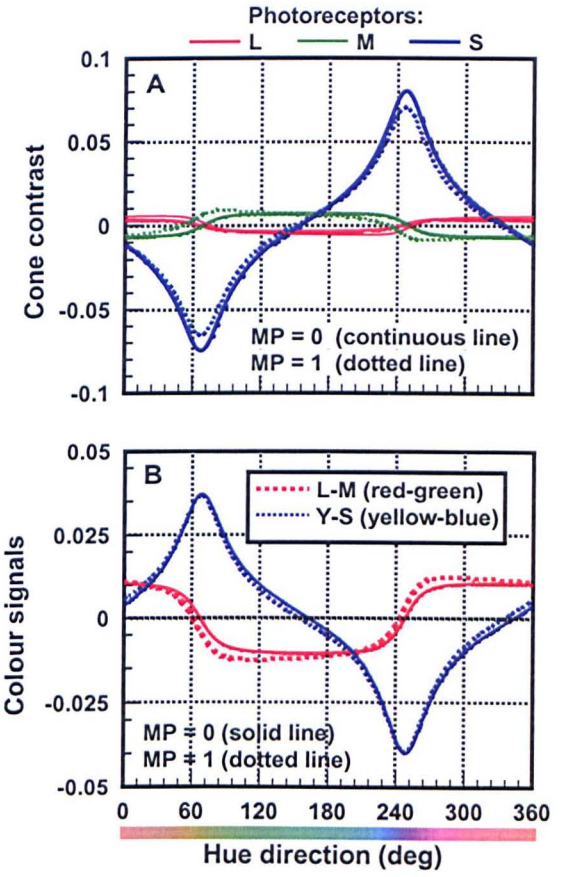


Figure 6.9: Analysis of cone contrasts and the corresponding colour signals generated for the mean colour detection thresholds shown as a black contour in Fig. 6.3A. (A) The solid lines plot the cone contrast signals generated in each class of cone photoreceptor for each direction in the CIE–(x, y) diagram. (B) The corresponding red-green and yellow-blue colour signals generated in the absence of MP. The attenuation of blue light by the MP can be simulated by filtering appropriately the light reaching the photoreceptors. The dotted lines plot the cone contrast signals (A) and the corresponding colour signals (B) when the light from the display is filtered by the MP. The results shown correspond to a MP filter with a peak optical density of 1 log unit (a value larger than that found in any subject investigated in this study).

6.5 Conclusions

The MAP test employs sinusoidal modulation, a small stimulus modulation depth and increments and decrements in mean luminance to facilitate the setting of the flicker-null condition. The technique is in good agreement with macular pigment data obtained on the Moreland anomaloscope. Ease of stimulus generation at any

required eccentricity makes the MAP test particularly suitable for measuring either 1D or 2D spatial profiles of MP optical density.

The spatial MPOD profiles measured after six months of supplementation with a combination of L (10 mg/day) and Z (10 mg/day) show a significant increase in MPOD, even at 6° eccentricity, making it essential to reference such measurements with respect to a larger eccentricity of 8° (Fig. 6.6C). Failure to do so underestimates uniformly the increase in MPOD as a result of supplementation.

Supplementation with a combination of L (10 mg/day) and Z (10 mg/day) over a period of six months increases MP distribution over ±8° around the macula, causing an almost uniform reduction in the percentage of transmitted blue light over the centre ±4° (Fig. 6.6D). Although a significant reduction in MP is observed in the supplemented group four months after supplementation was stopped, MPOD profiles in this group continue to remain higher than those in the placebo group.

For the stimulus conditions employed in this study YB chromatic detection thresholds do not differ significantly within the centre ±5° and do not correlate with the measured MPOD values, either at the fovea or in the periphery. All subjects showed high RG chromatic sensitivity (well within the normal range), but no correlation with MPOD values was found.

Hue cancellation techniques (Jameson and Hurvich, 1955) have also been employed to measure the spectral sensitivities of colour opponent mechanisms at a number of discrete locations in the visual field (Hibino, 1992). The author's findings suggest that the spectral sensitivity of the YB channel remains relatively constant with eccentricity in spite of large variation in MP optical density in the two subjects investigated. Although consistent with our findings, Hibino's experiments are more difficult to interpret because the test stimuli were scaled according to a single cortical magnification factor. Further, Hibino's findings do not require the gain change put forward since MP is not affected by YB chromatic sensitivity as shown in Fig. 6.9.

The cone contrast model developed to predict the effects of MP on colour detection thresholds accounts well for the observed experimental findings. In addition to predicting no changes in YB colour detection thresholds, even for peak MPOD values of one log unit, the model predicts a marginal improvement on RG colour thresholds with increased MP density. Contrary to expectations, the findings of this study suggest that at photopic levels of light adaptation increasing the density of the MP can only improve human chromatic discrimination sensitivity.

7 Summary and Conclusions

This thesis embodies a number of related studies of colour vision that involve accurate assessment of both 'normal' trichromats and colour deficient observers. The independent processing of luminance and chromatic signals, that forms much of the basis of this work, was previously demonstrated by Barbur et al. (1992) in pupillometric studies. The properties of the dynamic luminance contrast (LC) noise techniques developed to isolate the use of colour signals has been extended and investigated further. The parameters were optimised through a series of preliminary studies to ensure high sensitivity for detection of colour signals and the ability to classify normal trichromats and each class of colour deficiency; we named this the Colour Assessment and Diagnosis (CAD) test.

We compared the CAD test to a series of established clinical colour vision tests by examining a large group of colour deficient observers on all tests. Additionally, results in another large group of normal colour vision observers were used to evaluate the variability in chromatic sensitivity within normal trichromatism and to establish the normal limits on the CAD test. In another study, a comparison of chromatic detection thresholds and genetic results from DNA sequencing was undergone in order to explain some of the unusual colour deficiencies observed in different colour vision tests. Finally, CAD thresholds values for both the yellow-blue (YB) and red-green (RG) discrimination were compared with the macular

pigment (MP) density levels both at the fovea and in different regions of the retina.

Comparison of data from current tests

The most common measures of colour discrimination performance involve colour matching and colour detection thresholds. In Chapter 2 several colour vision tests that are normally used in a clinical environment have been described. Colour matching can be assessed on the Nagel anomaloscope, a compact clinical instrument that makes use of the Rayleigh match. A large group of subjects were assessed using Ishihara plates, AO-HRR plates, Farnsworth D15, City University (2^{nd} Ed.) test and the Nagel anomaloscope. Currently, the practice that provides an accurate description of colour vision loss involves the use of a battery of tests (Pokorny et al. 1979; Birch 2001). The outcome is a time consuming assessment that often yields inconsistent results. This difficulty has been object of concern within the Joint Aviation Authority (JAA) member states (Squire et al. 2005). The large variability and inconsistency in the results of colour vision tests suggest that an aspiring pilot may be rejected in one country and accepted in another. Chance guessing is often high which causes some observers that pass on one occasion to fail on re-test. The results of this investigation are in agreement with and extent of the findings of Squire's work. The analysis of results emphasise the importance of understanding the properties and the limitations of each test, if one is to be able to interpret the tests correctly. The comparison of the different clinical colour vision tests employed in this work reveal that only 55% of the subjects produce consistent results on the various tests. Currently, the Nagel anomaloscope is considered the reference test for grading and classifying the different types of deficiency (e.g. Birch 2001) and its parameters are often used to infer the severity of colour vision loss. The individual comparisons of the different colour vision tests versus the Nagel anomaloscope reveal, for example, that the Ishihara pseudoisochromatic plates have 100% sensitivity which makes this an excellent test for screening red/green colour deficient observers. However, it has also been shown that some observers with normal colour vision can fail the Ishihara plates

and that minimal colour deficient observers can show high within subject variability. The results of the AO-HRR test reveal that all normals are able to pass this test, but so do a small proportion of mild colour deficient observers. The results of this work have also shown that the number of errors on the Ishihara or AO-HRR plates does not correspond to the degree of colour vision loss. The ability to correctly classify the type of red/green colour deficiency is also limited in the Ishihara, AO-HRR, Farnsworth D15 and City University tests, as compared to the Nagel anomaloscope. Another important issue is that all these tests either do not address or are inefficient at assessing yellow/blue colour vision loss. Despite the very low number of congenital colour vision deficient observers that exhibit yellow/blue loss with this type of deficiency, most diseases of the retina and the optic nerve result in significant yellow/blue loss (e.g. optic neuritis, glaucoma). It would therefore be valuable to have a clinical colour vision test that screens accurately for both yellow/blue and red/green colour vision loss.

The CAD test

The development of the CAD test is described in Chapter 3. The CAD test involves a moving, colour-defined stimulus which is of the same mean luminance as the surrounding background and is buried in dynamic LC noise. Preliminary studies were carried out to optimise the parameters of the test and in particular the number of colour directions needed to achieve the desired results. Secondly, the parameters for background luminance and stimulus size were selected in order to achieve stable and sensitive measurement of chromatic sensitivity. This ensured that the measurement is robust against small variations of either size or background luminance. Sixteen colour directions are employed to isolate both YB and RG chromatic mechanisms; this number is larger than the minimum required, and has been selected to improve reliability and the detection of smaller changes in chromatic sensitivity. The results are graphed on an internationally recognised system of colour specification. Colour discrimination thresholds are plotted in the 1931 CIE-x,y chromaticity space in a manner similar to MacAdam's original experiments (MacAdam 1942). The elongations along the colour confusion axes

indicate the extent of colour vision loss. The results have shown that the CAD test classifies even minimal deficiencies into protan/deutan types in 99% of subjects with respect to the Nagel anomaloscope. It is however difficult to establish whether the misclassified subjects are indeed protanomalous, deuteranomalous or both. Further, repeated measurements available in 33 observers verified the repeatability of the CAD test revealing a strong correlation, 0.98 and 0.81, between the repeated measures for both RG and YB chromatic thresholds, respectively.

The 'standard' normal CAD observer

Chapter 4 has been devoted to establishing the statistical limits of the normal observer on the CAD test. A random population sample of 245 (128 males and 117 females, age range 14 to 60 yrs) observers were examined. Eight of these subjects produced RG thresholds that were clearly separated from the cluster of normal trichromats and were therefore not included in the analysis. The colour thresholds measured in the remaining 237 normal trichromats were used to calculate the 2.5% and 97.5% limits and the mean threshold values for RG and YB discrimination. These findings reveal the large variability and the asymmetry in the distribution of colour thresholds within normal trichromats. The existence of several subgroups within the normal population sample reflects the possible polymorphism in the L-cone amongst males with normal colour (Neitz & Jacobs 1986a; Neitz et al. 1995). Another explanation within the population of normal observers, which might explain the non-uniformity, is the possible effect on chromatic thresholds of heterozygous women (Jordan & Mollon 1993a). In addition to the group of normal trichromats, a large group of colour deficient observers (introduced in Chapter 3) were also analysed to validate the CAD test against the Nagel anomaloscope. The results of this study yield an agreement between the two tests of the order of 0.97. A study by Seshadri et al. (2005) showed that the coefficient of agreement between the Nagel anomaloscope and a web-based, uncalibrated version of the CAD test was 0.93. The CAD test was also showed to have a high screening efficiency with respect to the Nagel with a

specificity of 100% and a sensitivity of 97.5%. Results based on the CAD test pass/fail criteria for the normal observer were compared to the Nagel, Ishihara, AO-HRR, Farnsworth D15 and City University tests. The inconsistency in the overall results on these different colour vision tests was high (i.e. 45% of the subjects showed inconsistent results).

The genetics of anomalous observers

23 subjects with unusual colour deficiencies were selected in an attempt to explain the atypical and diverse results obtained in these subjects. The work in Chapter 5 describes a study relating the subject's genotype with the observed colour vision loss as evaluated by the Nagel and CAD tests. Genetic sequencing of photopigments genes was carried out in these subjects selected to represent both extremes of anomalous trichromacy. The results revealed differences in chromatic thresholds within a group that is undifferentiated with respect to their cone pigment genes. Secondly, the results show how phenotypically classified simple protanomalous observers can have very different pigment genes. The results of the CAD test for these subjects are explored in relation to their genotype. The major findings have shown that the RG thresholds assessed on the CAD test shows a clear dependence on the peak separation between the L- and the M-cone. However, in order to account more fully for the variety of colour discrimination loss, other factors such as changes in the optical density of the photopigments and post-receptoral amplification must be considered. From the genetics and CAD data we were able to infer that photopigment peak separations greater than 20 nm should be sufficient to convey normal trichromatic vision as identified by clinical colour vision tests.

Comparison of CAD thresholds and macular pigment

We also examined the effect of MP on chromatic sensitivity in a small group of subjects participating in a carotenoid supplementation trial. Previous studies have

shown that individual variations in macular pigment optical density (MPOD) can be as large as one log unit (Ruddock 1963). In Chapter 6, the objective of the study was to assess whether the variability in YB colour discrimination sensitivity, found in normal trichromatic subjects, can be attributed to differences in the pre-receptoral filtering of MPOD, both at the fovea and in the paracentral visual field. The measurement of spatial distribution of the MPOD revealed that as a result of supplementation with lutein and/or zeaxanthin, a significant increase in MPOD in the eye was found, particularly within a central region around the fovea subtending ±4°, but also even up to 6° eccentricity. Therefore, it is important to reference such measurements with respect to a larger eccentricity of ~8°. Failure to do so can lead to an overall underestimation of the increase in MPOD as a result of supplementation. The experimental data show no correlation between MP optical densities and colour thresholds, despite the known effect MP absorption of blue light has on colour matching experiments. Both YB and RG colour discrimination thresholds remain completely unaffected by the MP. The threshold colour discrimination model developed accounts well for the absence of any correlation between MP density and YB thresholds and predicts a marginal improvement in RG colour discrimination when the MP density is high.

7.1 Concluding remarks

The data obtained in this study show that the optimised CAD test fulfils the key requirements needed for an accurate colour vision assessment. Firstly, the spatiotemporal luminance contrast noise techniques employed isolate the use of colour signals. Secondly, the choice of the background level, stimulus size and the distribution of the 16 directions in 1931 CIE-x,y colour space makes the test robust and sensitive and provides the information needed to detect very small changes in chromatic sensitivity in repeated tests. The high correlation on re-testing is an indication of the high signal to noise ratio this test achieves (within each subject). These results reflect both the properties of fundamental RG and YB chromatic mechanisms and the measurement accuracy of the CAD technique. Thirdly, the establishment of the normal CAD observer provides a template that

allows instant detection and classification of even minimal deficiencies. The statistical distribution of results within normal trichromats can be used to investigate gender differences and give supporting evidence for interpretation of differences in genotype of the underlying photopigments. By establishing the statistical limits of colour discrimination in 'normal' trichromats, the CAD test has adequate sensitivity to detect and classify 'minimal' colour vision deficiencies that often produces variable results on repeated tests with conventional colour vision tests. In addition, the CAD test has enough specificity by minimising measurement variance within normal trichromats, to correctly identify normal colour vision.

The findings of this study, made possible by the sensitive and accurate measurements of chromatic sensitivity using the CAD test, also have important fundamental implications. The variability observed within normal trichromatism suggests the need for further studies to explain the existence of several subgroups with different mean RG chromatic thresholds. Additional molecular genetic analysis performed on subjects within a specific subgroup could provide further evidence for differences found in normal trichromatism. The results have revealed that the peak separation between pigments is not enough to fully account for the observed phenotypical differences. There are other factors such as optical density and differences in post-receptoral amplification of cone signals that act upon the photopigment λ_{max} differences and influence chromatic discrimination.

The large variability found in YB chromatic thresholds could have been explained by suggesting a correlation with the amount of MP in the eye. Higher and more variable YB thresholds in the fovea and the somewhat reduced variability measured in the near periphery where MPOD is reduced, might have suggested a direct relation to the amount of MP present in the retina. However, the results of this study failed to demonstrate a correlation between the MP density levels and chromatic discrimination thresholds, either in the fovea or in the paracentral retina. The model described in this study accounts for the lack of correlation between YB discrimination and MP and shows a slight gain in RG discrimination as a result of increased MP levels. The model predicts little or no change in YB colour discrimination, largely because the MP filter in front of the eye absorbs

blue light from both the background and the test pattern in a similar way. The relative independence of YB thresholds from absorption of blue light by the MP may not however extend to relatively low light levels when the signal to noise ratio in S-cone becomes small.

Trichromatic vision can provide us with functional advantages that can enhance our visual performance; these are often employed in various professional occupational tasks. It is therefore important to examine colour vision efficiently in groups of people with professions where normal colour vision or a certain level of colour discrimination is essential for safety reasons. The measurement of chromatic sensitivity on the CAD test could provide the link needed to set the minimum, safe colour discrimination sensitivity the subjects must have in occupations where colour vision is a critical task. It is also known that colour vision and particularly chromatic sensitivity, is a very good indicator of the normal functioning of the retina. Chromatic sensitivity is most affected in a number of diseases of the retina and the optic nerve. Accurate assessment of chromatic sensitivity can therefore assist in the detection of significant changes in chromatic sensitivity when monitoring the progress of a disease or a treatment.

Put together, the outcome of this study represents an important contribution towards a better understanding of both 'normal' and 'deficient' colour vision. In addition, these findings point the way towards how future research directions could be used to elucidate some remaining problems in understanding unusual colour vision deficiencies and the variability within 'normal' trichromats.

Appendices

Appendix A. Ishihara pseudoisochromatic plates.

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

Ishihara test		Nagel: Pass	Nagel: Fail	
Standard version 38-plate	Ishihara: Pass	22	1	κ=0.94 Specificity=91.7% (95% CI: 0.72-0.99)
1-25 plates (up to 3 misreadings)	Ishihara: Fail	2	155	Sensitivity=99.4% (95% CI: 0.96-1) 180 subjects

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

Ishihara test		Nagel: Pass	Nagel: Fail	
Standard version 38-plate	Ishihara: Pass	21	0	κ=0.93 Specificity=87.5% (95% CI: 0.67-0.97)
1-16 plates (no misreadings)	Ishihara: Fail	3	156	Sensitivity=100% (95% CI: 0.97-1) 180 subjects

Nagel Anomaloscope
Midpoint 36-44
Range 1-7

Ishihara test		Nagel: Pass	Nagel: Fail	
Standard version 38-plate	Ishihara: Pass	23	2	κ=0.94 Specificity=95.8% (95% CI: 0.77-1)
1-16 plates (up to 3 misreadings)	Ishihara: Fail	1	154	Sensitivity=98.7% (95% CI: 0.95-1) 180 subjects

Table A 1: Contingency tables comparing the Ishihara plate test pass/fail results with the Nagel anomaloscope. Failure on the Nagel anomaloscope was a matching range >7 and a midpoint that lies outside the RG scale 36-44. Failure on the Ishihara plates is as shown above (for a description of 'misreadings' see text).

A table showing which plates can be 'misread' can be found in Appendix D.

For the Ishihara pseudoisochromatic plate test the consensus for a pass in this study was that no errors and no misreadings are allowed on any plate. It is common practice, however, to allow up to three misreadings in certain plates (Table A.1a). Following Birch's recommendations (Birch 1997c) the criteria for a pass or a fail on the Ishihara test as a very efficient screening test (not intended for occupational colour vision screening) is to consider only the Transformation and Vanishing plates (first 16 plates) allowing up to a maximum of three misreadings but no errors. If we were to use this criterion in our study (Table A.1c), 98.7% of people with red/green colour deficiency would be identified, and the remaining 1.3% (all deuteranomalous trichromats) would be misclassed as having normal colour vision. These subjects were considered to have minimal colour deficiency, passing the D15 test and making one error only on the AO-HRR plates. This criterion would also suggest that 4.2% of people with normal colour vision would be misclassed as colour deficient by making errors that are not considered misreadings.

Appendix B. AO-HRR plates

AO-HRR (1 error on screening plates)		Nagel Anomaloscope Midpoint 36-44 Range 1-7: Pass	Nagel Anomaloscope Midpoint 36-44 Range 1-7: Fail	
	Pass	24	24	κ=0.67 Specificity=100% (95% CI: 0.83-1)
	Fail	0	130	Sensitivity=84.4% (95% CI: 0.77-0.90) 178 subjects

Table B 1: Contingency tables comparing the AO-HRR plate test pass/fail results with the Nagel anomaloscope. Failure on the Nagel anomaloscope was a matching range >7 and a midpoint that lies outside the RG scale 36-44. Failure on the AO-HRR plates is one error in the first 4 screening plates.

The criterion chosen for the AO-HRR in our study was based on no errors in any of the 20 plates. Previous studies have defined their pass/fail limit allowing for one error on the initial screening plates which identifies 96.4% of colour deficient observers (Birch 1997a), and have shown that 10-20% of normal trichromats may be expected to fail at least one screening plate, suggesting a significant overlap between the responses of normal and colour deficient people. By allowing no errors at all 94.2 % of colour deficients are correctly identified (see Fig. 2.3B), suggesting that 5.8% of people with abnormal colour vision can pass this test. If the criteria suggested by Birch were chosen, the sensitivity would reduce to 84.4% for this sample group and the sensitivity would remain 100% (Table B.1). The pass/limit that gave the highest overall efficiency for the observers tested was the one adopted in this study.

Appendix C. Mathematical definition of the parameters used in the analysis of colour vision tests.

Let $t \in \{c, \bar{c}\}$ denote the true condition of a subject, with c standing for normal colour vision and $\bar{c}$ standing for colour vision deficiency, and let $d \in \{p, \bar{p}\}$ denote the diagnosis of a subject from a certain colour vision test, with p meaning that the subject passed the test and $\bar{p}$ meaning that the subject failed the test. Then the definitions of the screening efficiency parameters become:

$$\begin{aligned} sensitivity &= P(d = \bar{p} \mid t = \bar{c}) \\ specificity &= P(d = p \mid t = c) \\ PPV &= P(t = \bar{c} \mid d = \bar{p}) \\ NPV &= P(t = c \mid d = p). \end{aligned} \qquad \text{(C.1)}$$

If a sample of individuals was tested both on the gold standard reference test (the Nagel anomaloscope for colour vision) and another test which is to be evaluated, then the results can be summarized in a 2x2 frequency table as follows:

		Reference test result		
		Pass	**Fail**	**Total**
Test	**Pass**	a	b	$a+b$
	Fail	c	d	$c+d$
	Total	$a+c$	$b+d$	N

Table C 1: General representation of a diagnostic test. Reproduced from (Altman 1991).

For example, a denotes the number of subjects who passed both the Nagel and the test that is evaluated, while b gives the number of subjects who failed the Nagel

and passed the other test. Then the screening efficiency parameters can be calculated as follows:

$$\begin{aligned} &sensitivity = d/(b+d) \\ &specificity = a/(a+c) \\ &PPV = d/(c+d) \\ &NPV = a/(a+b). \end{aligned} \qquad \textbf{(C.2)}$$

The name for the measure of agreement is kappa, written κ. It has a maximum of 1.00 when agreement is perfect; a value of zero indicates no agreement better than chance. This measure is defined as follows:

$$\kappa = \frac{p_o - p_e}{1 - p_e} \qquad \textbf{(C.3)}$$

where p_o is the overall proportion of agreement and p_e is the expected proportion of agreement by chance (Kundel and Polansky 2003), and are defined as:

$$\begin{aligned} p_o &= \frac{a + d}{N} \\ p_e &= \frac{(c + d)(b + d)}{N^2} + \frac{(a + b)(a + c)}{N^2} \end{aligned} \qquad \textbf{(C.4)}$$

To calculate the predictive ability (PPV and NPV) of the test for population groups with other prevalences of disease, i.e. for the general population (Altman 1991):

$$\begin{aligned} PPV &= \frac{\text{sen} \times \text{prevalence}}{\text{sen} \times \text{prevalence} + [(1 - \text{spec}) \times (1 - \text{prevalence})]} \\ NPV &= \frac{\text{spec} \times (1 - \text{prevalence})}{(1 - \text{sen}) \times \text{prevalence} + [\text{spec} \times (1 - \text{prevalence})]} \end{aligned}$$

where *sen* and *spec* stand for sensitivity and specificity, respectively.

Appendix D. Misreadings

A misreading is where a person fills in the 'partial loops' when looking at a number. An error is where a different number is seen on the transformation plates or no number at all is seen on the vanishing plates.

Plate no.	Normal Colour Vision	Misreading	Error by Colour deficient observer
1	12	12	12
2	8		3
3	6	8	5
4	29	28	70
5	57		35
6	5	8	2
7	3	8	5
8	15	18	17
9	74		21
10	2		-
11	6	8	-
12	97	87	-
13	45	46	-
14	5		-
15	7		-
16	16	18	-
17	73	75/78/13/23	-
18	-		See a number
19	-		See a number
20	-		See a number
21	-		See a number
22	26		Protans see the figure on the right only; deutans see the left figure only.
23	42		
24	35		
25	96		

Table D 1: Table of misreadings. Taken from the London Underground Occupational Health guidelines.

Appendix E. Statistical analysis for MPOD levels in the fovea and periphery for all groups.

Statistical analysis for MPOD levels in the fovea

Means Table for MPOD(0)
Effect: group

	Count	Mean	Std. Dev.	Std. Err.
C	5	.528	.080	.036
L	3	.522	.049	.028
P	6	.415	.121	.050
P*	1	.428	•	•
P-C	5	.535	.039	.017
Z	5	.514	.052	.023

ANOVA Table for MPOD(0)

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
group	5	.060	.012	1.896	.1424	9.481	.509
Residual	19	.121	.006				

Fisher's PLSD for MPOD(0)
Effect: group
Significance Level: 5 %

	Mean Diff.	Crit. Diff.	P-Value	
C, L	.006	.122	.9184	
C, P	.113	.101	.0302	S
C, P*	.099	.183	.2703	
C, P-C	-.007	.106	.8933	
C, Z	.014	.106	.7847	
L, P	.107	.118	.0728	
L, P*	.093	.193	.3245	
L, P-C	-.013	.122	.8270	
L, Z	.008	.122	.8931	
P, P*	-.014	.180	.8730	
P, P-C	-.120	.101	.0224	S
P, Z	-.099	.101	.0540	
P*, P-C	-.106	.183	.2397	
P*, Z	-.085	.183	.3416	
P-C, Z	.021	.106	.6842	

Table of measured mean MP optical density levels for each group. Results of the ANOVA test shows that the variance between groups is significant with a P-value of 0.14 compared to the variance within groups. Results of the Fisher's test for MPOD at the fovea for each group that received supplementation, show a significant increase in MP density for the C (p=0.03) and P-C (p=0.02) groups and slightly less for the Z (p=0.05) and L (p=0.07) groups in comparison to the placebo group. (Data supplied by W. Schalch).

Statistical analysis for MPOD levels in the periphery

Means Table for MPOD(2.5)
Effect: group

	Count	Mean	Std. Dev.	Std. Err.
C	5	.118	.028	.012
L	3	.114	.089	.051
P	6	.073	.053	.022
P*	1	.112	•	•
P-C	5	.180	.065	.029
Z	5	.107	.044	.020

ANOVA Table for MPOD(2.5)

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
group	5	.032	.006	2.126	.1065	10.629	.564
Residual	19	.058	.003				

Fisher's PLSD for MPOD(2.5)
Effect: group
Significance Level: 5 %

	Mean Diff.	Crit. Diff.	P-Value	
C, L	.004	.084	.9172	
C, P	.045	.070	.1896	
C, P*	.007	.126	.9112	
C, P-C	-.062	.073	.0912	
C, Z	.011	.073	.7491	
L, P	.041	.081	.3041	
L, P*	.003	.133	.9680	
L, P-C	-.066	.084	.1161	
L, Z	.007	.084	.8624	
P, P*	-.039	.124	.5247	
P, P-C	-.107	.070	.0045	S
P, Z	-.034	.070	.3199	
P*, P-C	-.069	.126	.2683	
P*, Z	.004	.126	.9415	
P-C, Z	.073	.073	.0489	S

Table of measured mean MP optical density levels for each group. Results of the ANOVA test shows that the variance between groups is significant with a P-value of 0.11 compared to the variance within groups. Results of the Fisher's test for MPOD in the periphery show a significant increase in MP density for the P-C (p=0.005) group in comparison to the placebo group and a significance difference between the P-C and Z groups (p=0.05). (Data supplied by W. Schalch).

Appendix F. 3D scatter plot of the Nagel anomaloscope matching parameters.

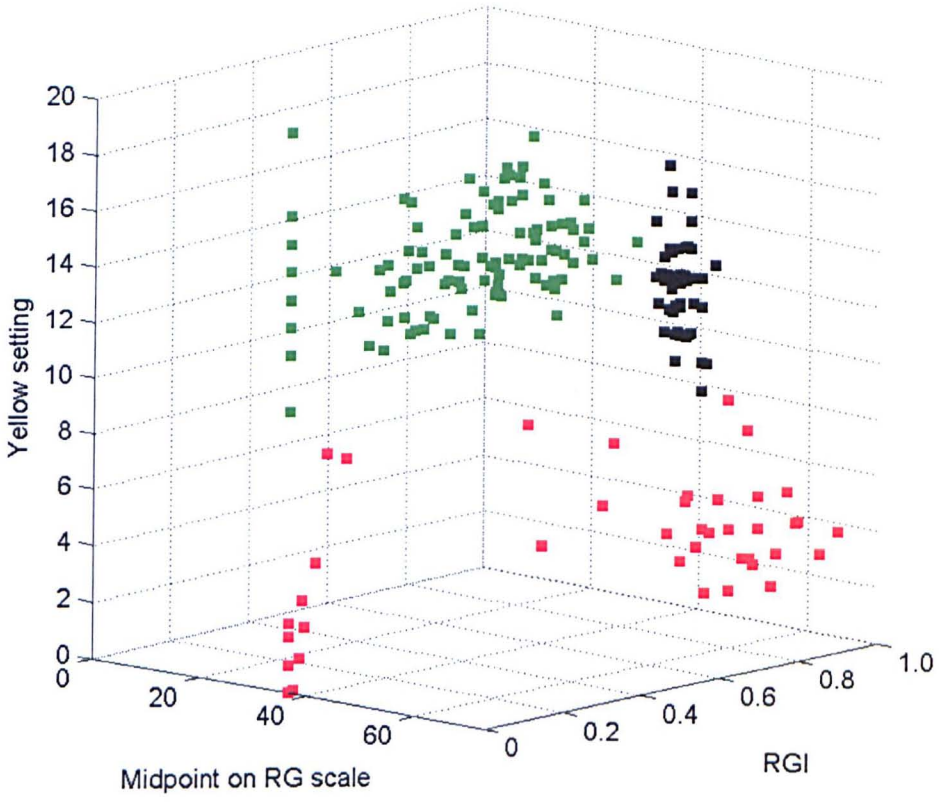


Figure F 1: 3D scatter plot of Nagel anomaloscope matching parameters for 231 observers that performed the test separated into deutan (green squares), protan (red squares) and normal (black squares) colour vision. The three orthogonal axes are: Nagel midpoint on the RG matching scale, RGI (range) and yellow setting required for the match. The yellow setting is recorded at both ends of the match. Here the value for the highest endpoint (towards the 'red' end) is plotted.

Fig. F 1 shows the same data used in Fig. 2.12 with an additional axis conveying information about the luminance setting needed to match the red-green mixture range. The orientation of the plot has been set to emphasise luminance differences in the red-green match of observers. The plot shows how deuteranopes and protanopes can be easily separated. Protan colour deficient observers usually have a relatively low luminance value at the red end of the red-green matching range, whilst deutan colour deficient observers tend to match both ends of the red-green range by equal amounts of luminance around 12-14. Normal colour vision observers have similar luminance settings to deutan observers.

References and bibliography

Abramov, I., Gordon, J., & Chan, H. (1991). Color appearance in the peripheral retina: effects of stimulus size. *Journal of the Optical Society of America, series A*, 8(2): 404-14.

Alpern, M., Lee, G. B., Maaseidvaag, F., & Miller, S. S. (1971). Colour vision in blue-cone 'monochromacy'. *Journal of Physiology*, 212(1): 211-33.

Alpern, M. & Pugh, E. N., Jr. (1977). Variation in the action spectrum of erythrolabe among deuteranopes. *Journal of Physiology*, 266(3): 613-46.

Alpern, M. (1979). Lack of uniformity in colour matching. *Journal of Physiology*, 288: 85-105.

Altman, D. G. (1991). Practical Statistics for Medical Research, 1st Ed., Chapman & Hall/CRC, London.

Asenjo, A. B., Rim, J., & Oprian, D. D. (1994). Molecular determinants of human red/green color discrimination. *Neuron*, 12(5): 1131-8.

Barbur, J. L. & Ruddock, K. H. (1980). Spatial characteristics of movement detection mechanisms in human vision. I. Achromatic vision. *Biological Cybernetics*, 37(2): 77-92.

Barbur, J. L., Holliday, I. E., & Ruddock, K. H. (1981). The spatial and temporal organisation of motion perception units in human vision. *Acta Psychologica*, 48(1-3): 35-47.

Barbur, J. L. & Forsyth, P. M. (1990). The effective contrast of coloured targets and its relation to visual search. In: Proceedings of the first international conference on visual search. D. Brogan (Ed.), London: Taylor & Francis Ltd.; 319-328.

Barbur, J. L. (1991). Pupillary responses to grating stimuli. *Journal of the Psychophysiology Society*, 5: 259-63.

Barbur, J. L. et al. (1992). Threshold and suprathreshold responses to chromatic stimuli using psychophysical and pupillometric methods. In: Non-invasive assessment of the visual system (Technical Digest Series). Vol.1. Washington DC.: Optical Society of America; 51-54.

Barbur, J. L., Birch, J., & Harlow, A. J. (1993). Colour vision testing using spatiotemporal luminance masking. In: Colour vision deficiencies XI. B.Drum (Ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands; 417-426.

Barbur, J. L., Harlow, A. J., & Plant, G. T. (1994). Insights into the different exploits of colour in the visual cortex. *Proceedings of the Royal Society of London B. Biological Sciences*, 258(1353): 327-34.

Barbur, J. L., Cole, V. A., & Plant, G. T. (1997). Chromatic discrimination in subjects with both congenital and acquired colour vision deficiencies. In: Colour vision deficiencies XIII, Documenta Ophthalmologica Proceedings Series 59. C.R. Cavonius (Ed.), Kluwer Academic Publishers, Dordrecht, Netherlands; 211-23.

Barbur, J. L. (2003). Understanding colour. *Trends in Cognitive Sciences*, 7(10): 434-6.

Barbur, J. L., De Cunha, D., Williams, C. B., & Plant, G. T. (2004). Study of instantaneous color constancy mechanisms in human vision. *Journal of Electronic Imaging*, 13(1): 15-28.

Barbur, J. L. (2004). 'Double-blindsight' revealed through the processing of color and luminance contrast defined motion signals. *Progress in Brain Research*, 144: 243-59.

Belcher, S. J., Greenshields, K. W., & Wright, W. D. (1958). Colour vision survey using the Ishihara, Dvorine, Bostrom and Kugelberg, Bostrom, and American-Optical Hardy-Rand-Rittler tests. *British Journal of Ophthalmology*, 42(6): 355-9.

Berendschot, T. T., Goldbohm, R. A., Klopping, W. A., van de Kraats J., van Norel, J., & Van Norren, D. (2000). Influence of lutein supplementation on macular pigment, assessed with two objective techniques. *Investigative Ophthalmology and Visual Science*, 41(11): 3322-6.

Berendschot, T. T. & Van Norren, D. (2006). *Investigative Ophthalmology and Visual Science*, 47(2): 709-14.

Birch, J. (1975). New diagnostic pseudoisochromatic plates for dichromats based on subtractive color matches. *American Journal of Optometry and Physiological Optics*, 52(6): 398-404.

Birch, J., Chisholm, I. A., Kinnear, P., Marre, M., Pinckers, A. J. L. G., Pokorny, J., Smith, V. C., & Verriest, G. (1979). Acquired color vision defects. In: Congenital and acquired color vision defects. J. Pokorny et al. (Eds.), Grune & Stratton, Inc., New York; 243-348.

Birch, J. (1989). Use of the Farnsworth-Munsell 100-hue test in the examination of congenital colour vision defects. *Ophthalmic and Physiological Optics*, 9(2): 156-62.

Birch, J., Barbur, J. L., & Harlow, A. J. (1992). New method based on random luminance masking for measuring isochromatic zones using high resolution colour displays. *Ophthalmic and Physiological Optics*, 12: 133-6.

Birch, J. & McKeever, L. M. (1993). Survey of the accuracy of new pseudoisochromatic plates. *Ophthalmic and Physiological Optics*, 13(1): 35-40.

Birch, J. (1997a). Clinical use of the American optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates for Red-green colour deficiency. *Ophthalmic and Physiological Optics*, 17(3): 248-54.

Birch, J. (1997b). Clinical use of the City University Test (2nd Edition). *Ophthalmic and Physiological Optics*, 17(6): 466-72.

Birch, J. (1997c). Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic and Physiological Optics*, 17(5): 403-8.

Birch, J. (2001). Diagnosis of defective colour vision, 2nd Ed., Butterworth-Heinemann, Oxford.

Birch, J. (2003). Extreme anomalous trichromatism. In: Normal and defective colour vision. J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), Oxford University Press; 364-369.

Bone, R. A., Landrum, J. T., & Tarsis, S. L. (1985). Preliminary identification of the human macular pigment. *Vision Research*, 25(11): 1531-5.

Bone, R. A., Landrum, J. T., Fernandez, L. & Tarsis, S. L. (1988). *Investigative Ophthalmology and Visual Science*, 29(6): 843-9.

Bone, R. A., Landrum, J. T., & Cains, A. (1992). Optical density spectra of the macular pigment in vivo and in vitro. *Vision Research*, 32(1): 105-10.

Bowmaker, J. K. & Dartnall, H. J. (1980). Visual pigments of rods and cones in a human retina. *Journal of Physiology*, 298: 501-11

Bowmaker, J. K. (1998). Visual pigments and molecular genetics of color blindness. *News in Physiological Sciences*, 13: 63-9.

Bowman, K. J. & Cole, B. L. (1980). A recommendation for illumination of the Farnsworth-Munsell 100-hue test. *American Journal of Optometry and Physiological Optics*, 57(11): 839-43.

Boynton, R. M., Schafer, W., & Neun, M. E. (1964). Hue-wavelength relation measured by color-naming method for three retinal locations. *Science*, 146: 666-8.

Brown, W. R. J. (1951). The influence of luminance level on visual sensitivity to color differences. *Journal of the Optical Society of America*, 41(10): 684-8.

Cambridge Research Systems Ltd. (2002). http://www.crsltd.com/apps/colour-vision/; Retrieved 18/11/2005.

Campbell, F. W. & Rushton, W. A. (1955). Measurement of the scotopic pigment in the living human eye. *Journal of Physiology*, 130(1): 131-47.

Casagrande, V. A. (1994). A third parallel visual pathway to primate area V1. *Trends in Neurosciences*, 17(7): 305-10.

CIE (2001a). Colours of light signals. Vienna, Austria: International Commission on Illumination. Standard: S004/E-2001.

CIE (2001b). International recommendations for color vision requirements for transport. International Commission on Illumination. CIE technical report, 143.

Ciulla, T. A., Curran-Celantano, J., Cooper, D. A., Hammond, B. R., Jr., Danis, R. P., Pratt, L. M., Riccardi, K. A., & Filloon, T. G. (2001). Macular pigment optical density in a midwestern sample. *Ophthalmology*, 108(4): 730-7.

Cole, B. L., Henry, G. H., & Nathan, J. (1966). Phenotypical variations of tritanopia. *Vision Research*, 6: 301-13.

Cole, B. L. (1963). Misuse of the Ishihara test for colour blindness. *British Journal of Physiological Optics*, 20: 113-8.

Conway, B. R. (2003). Colour vision: a clue to hue in V2. *Current Biology*, 13(8): R308-10.

Cornsweet, T. N. (1962). The staircase method in psychophysics. *American Journal of Psychology*, 75: 485-91.

Crognale, M. A., Teller, D. Y., Yamaguchi, T., Motulsky, A. G., & Deeb, S. S. (1999). Analysis of red/green color discrimination in subjects with a single X-linked photopigment gene. *Vision Research*, 39(4): 707-19.

Curcio, C. A., Allen, K. A., Sloan, K. R., Lerea, C. L., Hurley, J. B., Klock, I. B., & Milam, A. H. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *Journal of Comparative Neurology*, 312(4): 610-24.

Dacey, D. M. & Lee, B. B. (1994). The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature*, 367(6465): 731-5.

Dacey, D. M. (1996). Circuitry for color coding in the primate retina. *Proceedings of the National Academy of Sciences of the United States of America*, 93(2): 582-8.

Dacey, D. M. (1999). Primate retina:cell types, circuits and colour opponency. *Progress in Retinal and Eye Research*, 18(6): 737-63.

Dacey, D. M. & Packer, O. S. (2003). Colour coding in the primate retina: diverse cell types and cone-specific circuitry. *Current Opinion in Neurobiology*, 13(4): 421-7.

Dartnall, H. J., Bowmaker, J. K., & Mollon, J. D. (1983). Human visual pigments: microspectrophotometric results from the eyes of seven persons. *Proceedings of the Royal Society of London B. Biological Sciences*, 220(1218): 115-30.

Davies, N. P. & Morland, A. B. (2004). Macular pigments: their characteristics and putative role. *Progress in Retinal and Eye Research*, 23(5): 533-59.

De Monasterio, F. M., McCrane, E. P., Newlander, J. K., & Schein, S. J. (1985). Density profile of blue-sensitive cones along the horizontal meridian of macaque retina. *Investigative Ophthalmology and Visual Science*, 26(3): 289-302.

De Valois, R. L. & De Valois, K. K. (1993). A multi-stage color model. *Vision Research*, 33(8): 1053-65.

Derrington, A. M., Krauskopf, J. & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of Macaque, *Journal of Physiology*, 357: 241-65.

Deeb, S.S. (2004). Molecular genetics of colour vision deficiencies. *Clinical & Experimental Optometry: Journal of the Australian Optometrical Association*, 87(4-5): 224-29.

Farnsworth, D. (1943). The Farnsworth-Munsell 100 hue and dichotomous tests for colour vision. *Journal of the Optical Society of America*, 33: 568-78.

Farnsworth, D. (1947). The Farnsworth dichotomous test for color blindness - panel D-15. Psychological Corporation, New York.

Farnsworth, D. (1957). Testing for color deficiency in industry. *A. M. A. Archives of Industrial Health*, 16(2): 100-3.

Fitzgibbon, A., Pilu, M., & Fisher, R. B. (1999). Direct least square fitting of ellipses. *IEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5): 476-80.

Fletcher, R. (1972). Methods of examination of macular colour vision. A modified D-15 test. *Modern Problems in Ophthalmology*, 11: 22-4.

Franceschetti, A. (1928). Die Bedeutung der Einstellungstreite am Anomaloskop fur die Diagnose der einzelnen Typen der Farbensinnstorungen nebst Bermerkungen uber ihren Vererbungsmodus. *Schweizerische Medizinische Wochenschrift*, 52: 1273-8.

Frey, R. G. (1958). Most suitable pseudoisochromatic table for practice. Comparative studies on applicability of Bostrom, Bostrom-Kugelber, Hardy-Rand-Ritter, Ishihara, Rabkin and Stilling tables. *Albrecht von Graefe's Archiv für Ophthalmologie*, 160(3): 301-20.

Guild, J. (1931). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London*, 230A: 149-87.

Gunduz, K., Arden, G. B., & Perry, S. (1989). A new test of colour vision using TV and computer graphics: results in some common acquired eye diseases. In: Seeing colour and contour, J. J. Kulikowski, C. M. Dickinson, & I. J. Murray (Eds.), Pergamos Press, Oxford; 373-380.

Guth, S. L., Massof, R. W., & Benzschawel, T. (1980). Vector model for normal and dichromatic color vision. *Journal of the Optical Society of America*, 70(2): 197-212.

Hammond, B. R., Jr., Johnson, E. J., Russell, R. M., Krinsky, N. I., Yeum, K. J., Edwards, R. B., & Snodderly, D. M. (1997a). Dietary modification of human macular pigment density. *Investigative Ophthalmology and Visual Science*, 38(9): 1795-801.

Hammond, B. R., Jr., Wooten, B. R., & Snodderly, D. M. (1997b). Individual variations in the spatial profile of human macular pigment. *Journal of the Optical Society of America, series A*, 14(6): 1187-96.

Hammond, B. R., Jr. & Caruso-Avery, M. (2000). Macular pigment optical density in a Southwestern sample. *Investigative Ophthalmology and Visual Science*, 41(6): 1492-7.

Hammond, B. R., Jr., Ciulla, T. A., & Snodderly, D. M. (2002). Macular pigment density is reduced in obese subjects. *Investigative Ophthalmology and Visual Science*, 43(1): 47-50.

Handelman, G. J., Dratz, E. A., Reay, C. C., & van Kuijk, J. G. (1988). Carotenoids in the human macula and whole retina. *Investigative Ophthalmology and Visual Science*, 29(6): 850-5.

Hardy, L. H., Rand, G., & Rittler, M. C. (1954). The H.R.R polychromatic plates. *Journal of the Optical Society of America*, 44: 509-23.

Harris, E. K. & Boyd, J. C. (1995). Statistical bases of reference values in laboratory medicine. Marcel-Dekker, New York.

Hayashi, T., Motulsky, A. G., & Deeb, S. S. (1999). Position of a 'green-red' hybrid gene in the visual pigment array determines colour-vision phenotype. *Nature Genetics*, 22(1): 90-3.

He, J. C. & Shevell, S. K. (1994). Individual differences in cone photopigments of normal trichromats measured by dual Rayleigh-type color matches. *Vision Research*, 34(3): 367-76.

He, J. C. & Shevell, S. K. (1995). Variation in color matching and discrimination among deuteranomalous trichromats: theoretical implications of small differences in photopigments. *Vision Research*, 35(18): 2579-88.

Heard, P. F., Stone, C. J., Gregory, R. L., & Marmion, V. J. (1987). A new computer graphics test for red/green colour anomaly. In: Colour vision deficiencies VIII, Documenta Ophthalmologica Proceedings Series. G. Verriest (Ed.), Nijhoff-Dr. W. Junk, Dordrecht; 181-194.

Helve, J. (1972). A comparative study of several diagnostic tests of colour vision used for measuring types and degrees of congenital red-green defects. *Acta Ophthalmologica. Supplementum*, 115: 3-64.

Hill, A. R. (1987). How we see colour. In: Colour physics for industry. R. McDonald (Ed.), Society of Dyers and Colourists, England; 211-281.

Holmes, J. G. & Wright, W. D. (1982). A new colour perception lantern. *Color Research and Application*, 7: 82-8.

Hovis, J. K. & Neumann, P. (1995). Colorimetric analyses of various light sources for the D-15 color vision test. *Optometry and Vision Science*, 72(9): 667-78.

Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195(1): 215-43.

Hubel, D. H. & Wiesel, T. N. (1974). Sequence regularity and geometry of orientation columns in the monkey striate cortex. *Journal of Comparative Neurology*, 158(3): 267-93.

Hunt, R. W. G. (1991). Revised colour-appearance model for related and unrelated colours. *Color Research and Application*, 16: 146-65.

Hunt, R. W. G. (1998). Measuring Colour, 3rd Ed., Fountain Press, Kingston-upon-Thames, England.

Hurlbert, A. (2003). Colour vision: primary visual cortex shows its influence. *Current Biology*, 13(7): R270-72.

Hurvich, L. M. & Jameson, D. (1955). Some quantitative aspects of an opponent-colors theory. II Brightness, saturation, and hue in normal and dichromatic vision. *Journal of the Optical Society of America*, 45(8): 602-16.

Hurvich, L. M. & Jameson, D. (1956). Theoretical analysis of anomalous trichromatic color vision. *Journal of the Optical Society of America*, 46(12): 1075-89.

Hurvich, L. M. (1972). Colour vision deficiencies. In: Handbook of Sensory Physiology, vol. VII/4: Visual Psychophysics. D. Jameson & L. M. Hurvich (Eds.), Springer-Verlag, Berlin; 582-624.

Hurvich, L. M. (1981). Color Vision, Sinauer Associates, Sunderland, Mass., U.S.A.

Ingling, C. R., Jr. & Martinez-Uriegas, E. (1983). The relationship between spectral sensitivity and spatial sensitivity for the primate r-g X-channel. *Vision Research*, 23(12): 1495-500.

Joint Aviation Authorities (2002). JAR-FCL 3. Flight crew licensing (medical), Section 1, subpart B. Class 1 medical requirements (amendment 2). Hoofddorp, Netherlands: JAA.

Jagla, W. M., Breitsprecher, T., Kucsera, I., Kovacs, G., Wissinger, B., Deeb, S. S., & Sharpe, L. T. (2003). Hybrid pigment genes, dichromacy, and anomalous trichromacy. In: Normal and defective colour vision, J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), Oxford University Press, Oxford; 307-317.

Jameson, D. & Hurvich, L. M. (1955). Some quantitative aspects of an opponent-colors theory. I Chromatic responses and spectral saturation. *Journal of the Optical Society of America*, 45(7): 546-52.

Jordan, G. & Mollon, J. D. (1993a). A study of women heterozygous for colour deficiencies. *Vision Research*, 33(11): 1495-508.

Jordan, G. & Mollon, J. D. (1993b). The Nagel anomaloscope and seasonal variation of colour vision. *Nature*, 363(6429): 546-9.

Judd, D. B. & Wyszecki, G. (1975). Color in business science and industry, 3 Ed., Wiley, New York.

Kaas, J. H., Huerta, M. F., Weber, J. T., & Harting, J. K. (1978). Patterns of retinal terminations and laminar organization of the lateral geniculate nucleus of primates. *Journal of Comparative Neurology*, 182(3): 517-53.

Khachik, F., Bernstein, P. S., & Garland, D. L. (1997). Identification of lutein and zeaxanthin oxidation products in human and monkey retinas. *Investigative Ophthalmology and Visual Science*, 38(9): 1802-11.

King-Smith, P. E., Chioran, G. M., Sellers, K. L., & Alvarez, S. L. (1989). Normal and deficient colour discrimination analysed by colour television. In: Colour vision: physiology and psychophysics. J. D. Mollon & L. T. Sharpe (Eds.), Academic Press, London; 167-172.

Kiper, D. C. (2003). Colour and form in early stages of cortical processing. *Journal of Physiology*, 548(2): 335.

Kirschfeld, K. (1982). Carotenoid pigments: their possible role in protecting against photooxidation in eyes and photoreceptor cells. *Proceedings of the Royal Society of London B. Biological Sciences*, 216(1202): 71-85.

Knight, R., Buck, S. L., Fowler, G. A., & Nguyen, A. (1998). Rods affect S-cone discrimination on the Farnsworth-Munsell 100-hue test. *Vision Research*, 38(21): 3477-81.

Knoblauch, K. (1987). Age and illumination effects in the Farnsworth-Munsell 100-hue test. *Applied Optics*, 26(8): 1441-8.

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

Köpcke, W., Schalch, W., & LUXEA-Study Group (2005). Changes in macular pigment optical density following repeated dosing with lutein, zeaxanthin, or their combination in healthy volunteers - results of the LUXEA-study. *Investigative Ophthalmology and Visual Science*, 46: E-abstract 1768.

Kraft, T. W., Neitz, J., & Neitz, M. (1998). Spectra of human L cones. *Vision Research*, 38(23): 3663-70.

Kremers, J., Scholl, H. P., Knau, H., Berendschot, T. T., Usui, T., & Sharpe, L. T. (2000). L/M cone ratios in human trichromats assessed by psychophysics, electroretinography, and retinal densitometry. *Journal of the Optical Society of America, series A*, 17(3): 517-26.

Kundel, H. L. & Polansky, M. (2003). Measurement of observer agreement. *Radiology*, 228(2): 303-8.

Kvansakul J., Rodriguez-Carmona M., Edgar D.F., Barker F.M., Köpcke W., Schalch, W., & Barbur J.L. (2006). Supplementation with the carotenoids Lutein or Zeaxanthin improves visual performance. *Ophthalmic and Physiological Optics*, 26(4): 362-71.

Lakowski, R. (1958). Age and color vision. *Advances in Science*, 15: 231-6.

Lakowski, R. (1965a). Colorimetric and photometric data for the 10th edition of the Ishihara plates. *British Journal of Physiological Optics*, 22(4): 195-207.

Lakowski, R. (1965b). Testing of colour vision in prospective printers' apprentices and the problems the presents in selection. *British Journal of Physiological Optics*, 22: 10-32.

Lakowski, R. (1966) A critical evaluation of colour vision tests. Colour Group, GB.

Land, E. H. (1959). Experiments in color vision. *Scientific American*, 200(5): 84-94.

Landrum, J. T., Bone, R. A., Joa, H., Kilburn, M. D., Moore, L. L., & Sprague, K. E. (1997). A one year study of the macular pigment: the effect of 140 days of a lutein supplement. *Experimental Eye Research*, 65(1): 57-62.

Landrum, J. T., Bone, R. A., Moore, L. L., & Gomez, C. M. (1999). Analysis of zeaxanthin distribution within individual human retinas. *Methods in Enzymology*, 299: 457-67.

Lee, B. B., Martin, P. R., & Valberg, A. (1988). The physiological basis of heterochromatic flicker photometry demonstrated in the ganglion cells of the macaque retina. *Journal of Physiology*, 404: 323-47.

Lee, B. B. (2004). Paths to colour in the retina. *Clinical and Experimental Optometry*, 87(4-5): 239-48.

Lennie, P. & D'Zmura, M. (1988). Mechanisms of color vision. *Critical Reviews in Neurobiology*, 3(4): 333-400.

Lennie, P. (2003). Receptive fields. *Current Biology*, 13(6): R216-19.

Leventhal, A. G., Rodieck, R. W., & Dreher, B. (1981). Retinal ganglion cell classes in the Old World monkey: morphology and central projections. *Science*, 213(4512): 1139-42.

Ling, Y., & Hurlbert, A. C. (2005). Colour constancy is as good as colour memory allows – a new colour constancy index. *ECVP abstract.*

Livingstone, M. S. & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *The Journal of Neuroscience*, 4(1): 309-56.

MacAdam, D. L. (1942). Visual sensitivites to color differences in daylight. *Journal of the Optical Society of America, series A*, 32: 247-74.

MacLeod, D. I. (1972). Rods cancel cones in flicker. *Nature*, 235(5334): 173-4.

MacLeod, D. I. & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69(8): 1183-6.

Maxwell, J. C. (1855). Experiments on colours, as perceived by the eye, with remarks on colour-blindness. *Transactions of the Royal Society of Edinburgh*, 21: 275-98.

Merbs, S. L. & Nathans, J. (1992a). Absorption spectra of human cone pigments. *Nature*, 356(6368): 433-5.

Merbs, S. L. & Nathans, J. (1992b). Absorption spectra of the hybrid pigments responsible for anomalous color vision. *Science*, 258(5081): 464-6.

Michael, C. R. (1978). Color vision mechanisms in monkey striate cortex: dual-opponent cells with concentric receptive fields. *Journal of Neurophysiology*, 41(3): 572-88.

Mollon, J. D. (1989). "Tho' she kneel'd in that place where they grew..." The uses and origins of primate colour vision. *Journal of Experimental Biology*, 146: 21-38.

Mollon, J. D. & Bowmaker, J. K. (1992). The spatial arrangement of cones in the primate fovea. *Nature*, 360(6405): 677-9.

Moreland, J. D. & Kerr, J. (1978). Optimization of stimuli for trit-anomaloscopy. *Modern Problems in Ophthalmology*, 19: 162-6.

Moreland, J. D. & Kerr, J. (1979). Optimization of a Rayleigh-type equation for the detection of tritanomaly. *Vision Research*, 19(12): 1369-75.

Moreland, J. D. & Dain, S. L. (1993). Macular pigment contributes to variance in 100-hue tests. In: Colour vision deficiencies XII, Documenta Ophthalmologica Proceedings Series 57. B. Drum (Ed.), Kluwer Academic Publishers, Dordrecht, Netherlands; 517-22.

Moreland, J. D., Robson, A. G., & Kulikowski, J. J. (2001). Macular Pigment assessment using a colour monitor. *Color Research and Application*, S26: S261-S263.

Moreland, J. D. (2004). Macular pigment assessment by motion photometry. *Archives of Biochemistry and Biophysics*, 430(2): 143-8.

Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B., & Hogness, D. S. (1986a). Molecular genetics of inherited variation in human color vision. *Science*, 232(4747): 203-10.

Nathans, J., Thomas, D., & Hogness, D. S. (1986b). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science*, 232(4747): 193-202.

Nathans, J. (1999). The evolution and physiology of human color vision: insights from molecular genetic studies of visual pigments. *Neuron*, 24(2): 299-312.

Neitz, J. & Jacobs, G. H. (1986). Polymorphism of the long-wavelength cone in normal human colour vision. *Nature*, 323(6089): 623-5.

Neitz, J. & Jacobs, G. H. (1990). Polymorphism in normal human color vision and its mechanism. *Vision Research*, 30(4): 621-36.

Neitz, M., Neitz, J., & Jacobs, G. H. (1991). Spectral tuning of pigments underlying red-green color vision. *Science*, 252(5008): 971-4.

Neitz, J., Neitz, M., & Jacobs, G. H. (1993). More than three different cone pigments among people with normal color vision. *Vision Research*, 33(1): 117-22.

Neitz, M. & Neitz, J. (1995). Numbers and ratios of visual pigment genes for normal red-green color vision. *Science*, 267(5200): 1013-6.

Neitz, M., Neitz, J., & Grishok, A. (1995). Polymorphism in the number of genes encoding long-wavelength-sensitive cone pigments among males with normal color vision. *Vision Research*, 35(17): 2395-407.

Neitz, J., Neitz, M., & Kainz, P. M. (1996). Visual pigment gene structure and the severity of color vision defects. *Science*, 274(5288): 801-4.

Neitz, J., Neitz, M., He, J. C., & Shevell, S. K. (1999). Trichromatic color vision with only two spectrally distinct photopigments. *Nature Neuroscience*, 2(10): 884-8.

Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*, 17: 857-872.

Osorio, D. & Vorobyev, M. (1996). Colour vision as an adaptation to frugivory in primates. *Proceedings. Biological sciences/The Royal Society,* 263(1370): 593-9.

Oyster, C. (1999). The human eye: structure and function. Sinauer Associates Ltd., Massachusetts:Sunderland.

Pacheco-Cutillas, M., Edgar, D. F., & Sahraie, A. (1999). Acquired colour vision defects in glaucoma-their detection and clinical significance. *British Journal of Ophthalmology*, 83(12): 1396-402.

Pinckers, A. & Baron, J. (1978). Clinical evaluation of Lanthony's new color test and desaturated 15-hue. *Modern Problems in Ophthalmology*, 19: 144.

Pitt, F. H. G. (1944). Monochromatism. *Nature*, 154: 466-8.

Pitt, F. H. G. (1935). Characteristics of dichromatic vision. In: Committee on the physiology of vision, report no. 14, His Majesty's Stationary Office, London; 5-58.

Pokorny, J., Smith, V. C., & Swartley, R. (1970). Threshold measurements of spectral sensitivity in a blue monocone monochromat. *Investigative Ophthalmology*, 9(10): 807-13.

Pokorny, J., Smith, V. C., & Katz, I. (1973). Derivation of the photopigment absorption spectra in anomalous trichromats. *Journal of the Optical Society of America*, 63(2): 232-7.

Pokorny, J. & Smith, V. C. (1977). Evaluation of single-pigment shift model of anomalous trichromacy. *Journal of the Optical Society of America*, 67(9): 1196-209.

Pokorny, J., Smith, V. C., Verriest, G., & Pinckers, A. J. L. G. (1979). Congenital and Acquired Color Vision Defects, Grune & Stratton, New York.

Pokorny, J., Smith, V. C., & Went, L. N. (1981). Color matching in autosomal dominant tritan defect. *Journal of the Optical Society of America*, 71(11): 1327-34.

Pokorny, J., Smith, V., & Lutze, M. (1987). Aging of the human lens. *Applied Optics*, 26: 1437-40.

Pokorny, J. & Smith, V. C. (1997). How much light reaches the retina? In: Colour vision deficiencies XIII, C. R. Cavonius (Ed.), Kluwer Academic Publishers, Dordrecht; 491-511.

Polyak, S. L. (1941). The retina. University of Chicago Press, Chicago.

Reading, V. M. & Weale, R. A. (1974). Macular pigment and chromatic aberration. *Journal of the Optical Society of America*, 64(2): 231-4.

Reffin, J. P., Astell, S., & Mollon, J. D. (1989). Trials of a computer-controlled colour vision test that preserves the advantages of pseudoisochromatic plates. In: Colour vision deficiencies X. B. Drum, J. D. Moreland, & A. Serra (Eds.), Kluwer Academic Publishers, Dordrecht, Netherlands; 69-76.

Regan, B. C., Reffin, J. P., & Mollon, J. D. (1994). Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Research*, 34(10): 1279-99.

Robson, A. G., Moreland, J. D., Pauleikhoff, D., Morrissey, T., Holder, G. E., Fitzke, F. W., Bird, C., & van Kuijk, F. J. G. M. (2003). Macular pigment density and distribution: comparison of fundus autofluorescence with minimum motion photometry. *Vision Research*, 43(16): 1765-75.

Rodieck, R. W. (1991). Which cells code for colour? In: From pigments to perception: advances in understanding visual processes. A. Valberg & B. B. Lee, (Eds.), Plenum Press, New York; 83-93.

Rodieck, R. W. & Watanabe, M. (1993). Survey of the morphology of macaque retinal ganglion cells that project to the pretectum, superior colliculus, and parvicellular laminae of the lateral geniculate nucleus. *Journal of Comparative Neurology*, 338(2): 289-303.

Rodriguez-Carmona, M., Harlow, J. A., & Barbur, J. L. (2003). Pupil responses to chromatic afterimages. *Investigative Ophthalmology and Visual Science*, 44: E-abstract 1917.

Rodriguez-Carmona, M., Harlow, J. A., Grace, W. & Barbur, J. L. (2005). The variability of normal trichromatic vision and the establishment of the 'normal' range. In: Proceedings of the 10th Congress of the International Colour Association, Granada (Granada, 2005): 979-82

Rodriguez-Carmona, M., Kvansakul, J., Harlow, J. A., Köpcke, W., Schalch, W. & Barbur, J. L. (2006). The effects of supplementation with lutein and/or zeaxanthin on human macular pigment density and colour vision. *Ophthalmic and Physiological Optics*, 26(2): 137-47.

Roorda, A. & Williams, D. R. (1999). The arrangement of the three cone classes in the living human eye. *Nature*, 397(6719): 520-2.

Ruddock, K. H. (1963). Evidence for macular pigmentation from colour matching data. *Vision Research*, 3(9-10): 417-29.

Ruddock, K. H. (1965). The effect of age upon colour vision--II. changes with age in light transmission of the ocular media. *Vision Research*, 5(1-3): 47-58.

Ruddock, K. H. (1972). Observer variations in foveal colour vision responses. *Vision Research*, 12(1): 145-9.

Rushton, W. A. (1963). A cone pigment in the protanope. *Journal of Physiology*, 168: 345-59.

Rushton, W. A. (1966). Densitometry of pigments in rods and cones of normal and color defective subjects. *Investigative Ophthalmology*, 5(3): 233-41.

Rushton, W. A. (1972). Visual pigments in man. In: Handbook of Sensory Physiology, vol. VII/1: Photochemistry of Vision. H. J. A. Dartnall (Eds.), Springer-Verlag, New York; 364-94.

Schalch, W. (1992). Carotenoids in the retina--a review of their possible role in preventing or limiting damage caused by light and oxygen. *EXS*, 62: 280-98.

Schalch, W., Rodriguez-Carmona, M., Harlow, J. A., Barbur, J. L., & Köpcke, W. (2004). Macular pigment optical density (MPOD) measurements using visual displays - a new method and first results. *Investigative Ophthalmology and Visual Science*, 45: E-abstract 1296.

Schalch, W., Barker, F. M., & LUXEA-Study Group (2005). Ocular and general safety of supplementation with zeaxanthin and lutein; plasma exposure levels of carotenoids and 3'-dehydro-lutein - results of the LUXEA-study. *Investigative Ophthalmology and Visual Science*, 46: E-abstract 1765.

Schiller, P. H. & Lee, K. (1991). The role of the primate extrastriate area V4 in vision. *Science*, 251(4998): 1251-3.

Schmidt, I. (1955). Some problems related to testing color vision with the Nagel anomaloscope. *Journal of the Optical Society of America*, 45: 515.

Schneeweis, D. M. & Schnapf, J. L. (1995). Photovoltage of rods and cones in the macaque retina. *Science*, 268(5213): 1053-6.

Schwartz, S. H. (2004). Visual Perception: a clinical orientation. 3 Ed., McGraw-Hill, New York.

Seshadri, J., Christensen, J., Lakshminarayanan, V., & Bassi, C. J. (2005). Evaluation of the new web-based "colour assessment and diagnosis" test. *Optometry and Vision Science*, 82(10): 882-5.

Sharpe, L. T. & Nordby, K. (1990). Total colour-blindness: an introduction. In: Night vision: basic, clinical and applied aspects. R. F. Hess, L. T. Sharpe, & K. Nordby (Eds.), Cambridge University Press, Cambridge; 253-289.

Sharpe, L. T., Stockman, A., Jagle, H., Knau, H., Klausen, G., Reitner, A., & Nathans, J. (1998a). Red, green, and red-green hybrid pigments in the human retina: correlations between deduced protein sequences and psychophysically measured spectral sensitivities. *Journal of Neuroscience*, 18(23): 10053-69.

Sharpe, L. T., Stockman, A., Knau, H., & Jagle, H. (1998b). Macular pigment densities derived from central and peripheral spectral sensitivity differences. *Vision Research*, 38(21): 3233-9.

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

Sherman, S. M. (1996). Dual response modes in lateral geniculate neurons: mechanisms and functions. *Vision Neuroscience*, 13(2): 205-13.

Shevell, S. K. & He, J. C. (1997a). The visual photopigments of simple deuteranomalous trichromats inferred from color matching. *Vision Research*, 37(9): 1115-27.

Shevell, S. K. & He, J. C. (1997b). Phenotypes of anomalous trichromacy. In: Colour vision deficiencies XIII. C.R.Cavonius (Ed.), Kluwer Academic Publishers, Dordrecht; 45-58.

Sloan, L. L. & Habel, A. (1956). Tests for color deficiency based on the pseudoisochromatic principle; a comparative study of several new tests. *A.M.A. Archives in Ophthalmology*, 55(2): 229-39.

Smith, V. C. & Pokorny, J. (1972). Spectral sensitivity of color-blind observers and the cone photopigments. *Vision Research*, 12(12): 2059-71.

Smith, V. C. & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15(2): 161-71.

Smith, V. C., van Everdingen, J., & Pokorny, J. (1991) Sensitivity of arrangement tests as evaluated in normals at reduced levels of illumination. In: Colour vision deficiencies X. Documenta Ophthalmologica Proceedings Series. B. Drum, J. D. Moreland, & A. Serra (Eds.), Kluwer, Dordrecht; 177-185.

Snodderly, D. M., Auran, J. D., & Delori, F. C. (1984a). The macular pigment. II. Spatial distribution in primate retinas. *Investigative Ophthalmology and Visual Science*, 25(6): 674-85.

Snodderly, D. M., Brown, P. K., Delori, F. C., & Auran, J. D. (1984b). The macular pigment. I. Absorbance spectra, localization, and discrimination from other yellow pigments in primate retinas. *Investigative Ophthalmology and Visual Science*, 25(6): 660-73.

Snodderly, D. M. (1995). Evidence for protection against age-related macular degeneration by carotenoids and antioxidant vitamins. *American Journal of Clinical Nutrition*, 62(6): 1448S.

Snodderly, D. M. (1999). Nutritional and Environmental influences on the eye. A. Taylor (Ed.), CRC Press: Boca Raton, Florida.

Sommerburg, O., Keunen, J. E. E., Bird, A. C., & van Kuijk, F. J. G. M. (1998). Fruits and vegetables that are sources for lutein and zeaxanthin: the macular pigment in human eyes. *British Journal of Ophthalmology*, 82(8): 907-10.

Sommerburg, O. G., Siems, W. G., Hurst, J. S., Lewis, J. W., Kliger, D. S., & van Kuijk, F. J. (1999). Lutein and zeaxanthin are associated with photoreceptors in the human retina. *Current Eye Research*, 19(6): 491-5.

Speranskaya, N. I. (1959). Determination of spectrum co-ordinates for twenty-seven normal observers. *Optics and Spectroscopy*, 7: 424.

Squire, T. J., Rodriguez-Carmona, M., Evans, A. D. B., & Barbur, J. L. (2005). Color vision tests for aviation: comparison of the anomaloscope and three lantern types. *Aviation, Space and Environmental Medicine*, 76: 421-9.

Steward, J. M. & Cole, B. L. (1989). What do color vision defectives say about everyday tasks? *Optometry and Vision Science*, 66(5): 288-95.

Stiles, W. S. & Burch, J. M. (1959). N.P.L. colour-matching investigation: final report (1958). *Optica Acta*, 6: 1.

Stockman, A. & Sharpe, L. T. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13): 1711-37.

Stockman, A., Sharpe, L. T., & Fach, C. (1999). The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches. *Vision Research*, 39(17): 2901-27.

Svaetichin, G. (1953). The cone action potential. *Acta Physiologica Scandinavica*, 29, 565-599.

Svaetichin, G. (1958). Component analysis of action potentials from single neurons. *Experimental Cell Research*, 14(S5): 234-61.

Trendelenburg, W. (1929). Zur Diagnostic des abnormen Farbensinnes. *Klinisches Monatsblatt für Augenheilkunde*, 83: 721-32.

Troxler, D. (1804). Uber das Verschwindern gegebener Gegenstande innerhalb unsers Gesichtskrcises. In: Ophthalmologisches Bibliothek, Vol. II., J. A. Schmidt & K. Himley (Eds.), Jena: Fromann; 51-53.

Van Heel, L., Went, L. N., & Van Norren, D. (1980). Frequency of tritan disturbances in a population study. In: Colour vision deficiencies V. G. Verriest (Ed.), Adam Hilger, Bristol; 256-260.

Van Norren, D. & Vos, J. J. (1974). Spectral transmission of the human ocular media. *Vision Research*, 14(11): 1237-44.

Van Norren, D. & Went, L. N. (1981). New test for the detection of tritan defects evaluated in two surveys. *Vision Research*, 21(8): 1303-6.

Ventura, D. F., Silveira, L. C. L., Rodrigues, A. R., De Souza, J. M., Gualtieri, M., Bonci, D., & Costa, M. F. (2003). Preliminary norms for the Cambridge colour test. In: Normal and defective colour vision. J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), Oxford University Press, Oxford; 331-339.

Verriest, G. (1963). Further studies on acquired deficiency of color discrimination. *Journal of the Optical Society of America*, 53: 185-95.

Verriest, G. (1968). Comparative study of effectiveness of various tests for recognition of color vision abnormalities. *Archives des maladies professionnelles de médecine du travail et de sécurité sociale*, 29(6): 293-314.

Vollrath, D., Nathans, J., & Davis, R. W. (1988). Tandem array of human visual pigment genes at Xq28. *Science*, 240(4859): 1669-72.

Walls, G. L. (1959). Peculiar color blindness in peculiar people. *A.M.A. Archives in Ophthalmology*, 62(1): 13-32.

Weale, R. A. (1953). Cone-monochromatism. *Journal of Physiology*, 121: 548-69.

Weale, R. A. (1988). Age and transmittance of the human crystalline lens. *Journal of Physiology*, 395: 577-87.

Weiter, J. J., Delori, F., & Dorey, C. K. (1988). Central sparing in annular macular degeneration. *American Journal of Ophthalmology*, 106(3): 286-92.

Weitz, C. J., Went, L. N., & Nathans, J. (1992). Human tritanopia associated with a third amino acid substitution in the blue-sensitive visual pigment. *American Journal of Human Genetics*, 51(2): 444-6.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22: 209-212.

Winderickx, J., Lindsey, D. T., Sanocki, E., Teller, D. Y., Motulsky, A. G., & Deeb, S. S. (1992). Polymorphism in red photopigment underlies variation in colour matching. *Nature*, 356(6368): 431-3.

Wolffsohn, J. S., Cochrane, A. L., Khoo, H., Yoshimitsu, Y., & Wu, S. (2000). Contrast is enhanced by yellow lenses because of selective reduction of short-wavelength light. *Optometry and Vision Science*, 77(2): 73-81.

Wooten, B. R., Hammond, B. R., Jr., Land, R. I., & Snodderly, D. M. (1999). A practical method for measuring macular pigment optical density. *Investigative Ophthalmology and Visual Science*, 40(11): 2481-9.

Wooten, B. R. & Hammond, B. R. (2002). Macular pigment: influences on visual acuity and visibility. *Progress in Retinal and Eye Research*, 21(2): 225-40.

Wright, W. D. (1928, 1929). A re-determination of the mixture curves of the spectrum. *Transactions of the Optical Society of London*, (30): 141-64.

Wright W.D. & Pitt, F. H. G. (1934). Hue discrimination in normal colour-vision. *Proceedings of the Physiological Society of London*, 46: 459-473.

Wright W.D. (1941). The sensitivity of the eye to small colour differences. *Proceedings of the Physiological Society of London*, 53: 93-112.

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

Wright, W. D. (1952). The characteristics of tritanopia. *Journal of the Optical Society of America*, 42(8): 509-21.

Wyszecki, G. & Stiles, W. S. (1982). Color Science: concepts and methods, quantitative data and formulae. 2 Ed., John Wiley & Sons, Inc., New York.

Yeh, T., Pokorny, J., & Smith, V. C. (1993). S-cone discrimination sensitivity and performance on arrangement tests. In: Colour vision deficiencies XI, Documenta Ophthalmologica Proceedings Series, B. Drum (Ed.), Kluwer, Dordrecht; 293-302.

Young, R. W. (1978). Visual cells, daily rhythms, and vision research. *Vision Research*, 18(5): 573-8.

Young, R. S. & Alpern, M. (1980). Pupil responses to foveal exchange of monochromatic lights. *Journal of the Optical Society of America*, 70(6): 697-706.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology*, 236(3): 549-73.

Zeki, S. M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology*, 277: 273-90.

Zeki, S. (1980). The representation of colours in the cerebral cortex. *Nature*, 284(5755): 412-8.

Zeki, S. (1993). A vision of the brain. Blackwell Scientific Publications, Oxford.

Zrenner, E. (1985) A new concept for the contribution of retinal colour-opponent ganglion cells to hue discrimination and colour constancy: the zero signal detector. In: Central and peripheral mechanisms of colour vision: proceedings of an international symposium held at the Wenner Gren Center Stockholm, June 14-15, 1984. D. Ottoson & S. Zeki (Eds.), Macmillan, London; 165-181.