The asymmetries of colour constancy as determined through illumination discrimination using tuneable LED light sources

Bradley Pearce

Institute of Neuroscience

Faculty of Medical Sciences, Newcastle University

Newcastle upon Tyne, NE2 4HH.

Thesis submitted for the degree of Doctor of Philosophy

April 2015

Typeset in Century, by Morris Fuller Benton.

Preface

A list of publications, for this and related projects can be seen listed below:

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G., & Hurlbert, A. (2014). Chromatic Illumination Discrimination Reveals that Human Colour Constancy is Optimised for Blue Daylight Illuminations. PLoS one 9(2), e87989.

Pearce, B., Radonjic, A., Dubin, H., Cottaris, N. P., Mackiewicz, M., Finlayson, G., Brainard, D. H., & Hurlbert, A. (2014). Illumination discrimination reveals "blue" bias for colour constancy in real and simulated scenes. Journal of Vision, 14(10). doi: 10.1167/14.10.599.

Hurlbert, A., **Pearce, B.**, Mackiewicz, M., & Finlayson, G. (2014). Colour Constancy in Immersive Viewing. Journal of Vision, 14(10). doi: 10.1167/14.10.794

Finlayson, G., Mackiewicz, M., Hurlbert, A., **Pearce, B.**, & Crichton, S. (2014). On calculating metamer sets for spectrally tunable LED illuminators. JOSA A 31(7), 1577-1587.

Pearce, B., Mackiewicz, M., Crichton, S., Finlayson, G., & Hurlbert, A. (2013). Are colour constancy mechanisms biased for typical illuminations? AIC 2013.

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G., & Hurlbert A. (2012). Colour constancy by illumination matching in real world scenes. Journal of Vision 12(9), 57.

Mackiewicz, M., **Pearce, B.**, Crichton, S., Finlayson, G., & Hurlbert, A (2012) .Achromatic Adjustment outdoors using MEMS reflective display. Perception 41(12), 1522.

Crichton, S., **Pearce, B.**, Mackiewicz, M., Finlayson, G., & Hurlbert, A. (2012). The illumination correction bias of the human visual system. Journal of Vision 12(9), 64.

Krieger, A., Dubin, H., **Pearce, B.**, Aston, A., Hurlbert, A. C., Brainard, D. H., & Radonjic, A. (in press). Illumination discrimination depends on scene surface ensemble. Vision Sciences Society, Annual Meeting, 2015.

Cranwell, M. B., **Pearce, B.**, Loveridge, C., & Hurlbert, A. C. (in press). Performance on the Farnsworth-Munsell 100-Hue test is significantly related to non-verbal IQ. Investigative Ophthalmology and Visual Science.

Abstract

The light reflected from object surfaces changes with the spectral content of the illumination. Despite these changes, the human visual system tends to keep the colours of surfaces constant, a phenomenon known as *colour constancy*. Colour constancy is known to be imperfect under many conditions; however, it is unknown whether the underlying mechanisms present in the retina and the cortex are optimised for the illuminations under which they have evolved, namely, natural daylights, or for particular objects. A novel method of measuring colour constancy, by illumination discrimination, is presented and explored. This method, unlike previous methods of measuring colour constancy, allows the testing of multiple, real, illuminations with arbitrary spectral content, through the application of tuneable, multi-channel LED light sources. Data from both real scenes, under real illuminations, and computer simulations are presented which support the hypothesis that the visual system maintains higher levels of colour constancy for daylight illumination changes, and in particular in the "bluer" direction, which are also the changes most frequent in nature. The low-level cone inputs for various experimental scenes are examined which challenge all traditional theories of colour constancy supporting the conclusions that higher-level mechanisms of colour constancy are biased for particular illuminations. Furthermore, real and simulated neutral (grey) surfaces are shown to affect levels of colour constancy. Moreover, the conceptual framework for discussing colour constancy with respect to emergent LED light sources is discussed.

Acknowledgements

I would firstly like to thank the EPSRC for supporting me through this studentship (Grant ref: EP/H022325/1). I have had the opportunity to meet great people while conducting this work. These people have sharpened, and tempered my mind, and for this I thank them. In particular, Anya Hurlbert, my supervisor and dear friend, whose contribution to my mind cannot be summarised in a few words, and whose intellect is only surpassed by her kindness and integrity. I thank Max Hammerton, for inspiring me on many fronts, for his dear friendship, and for his wisdom, which I have not yet exhausted in 3 years of lunchtime conversation. I also thank Stacey Aston, Matthew Cranwell and Stuart Crichton, who have taught me the meaning of camaraderie, and who are true friends.

The experiences and values I most cherish I owe to my family: Auntie Gay, Uncle John, Gabe, Auntie Ann, Uncle Harry, Nanny Gul and Grandad Parker, Auntie Trace, Christian, Harry, India and most of all Mum; I love you all, what I am is because of you – to those that find my wit distasteful, please find a list of accomplices named above.

We are patterns in the fabric of the universe, living on a tiny speck of dust, within an ocean of forces, of cosmic scale. Despite this scale, the never-ending complexity and novelty of species on our planet, and their ability to exploit every evolutionary niche, is humbling. We, humans, have evolved the ability to specialise and trade labours; leading to a surplus of energy, advancement in technology and, for some, a crusade towards an understanding of that very fabric of which we are made. No such scientific endeavour can be undertaken without the aid of the entire community, who collectively support such specialisation; providing healthcare, transport, manufacturing food and so forth. This work, for whatever it is worth, belongs to the people, in payment for such cooperation.

Contents

Chapter 1: 1

An Introduction to Colour Constancy 1

Quantifying Colour 6

The Nature of the problem 9

Experimental Approaches to Colour Constancy. 15

Colour Constancy and the Ecological Hypothesis. 25

Objectives 28

Chapter 2: 30

Materials and Calibration Methods 30

Setup Overview 31

Viewing box 32

Tuneable LED Light Sources 33

Calibrating LED light sources 35

Methods 36

Apparatus 36

Design 36

Procedure 37

Results 38

Spectral and Colorimetric Fitting 41

Creating real surfaces with specific chromaticities 46

Methods 47
Design and Procedure 48
Analysis and Modelling 49
Chapter 3: 51
Colour constancy by illumination matching 51
Experiment 1.1 55
Methods 55
Ethics and declaration 55
Participants 55
Design 56
Apparatus 56
Stimuli 58
Procedure 60
Results 62
Discussion 64
Experiment 1.2 67
Methods 67
Apparatus and Materials 67
Procedure 67
Analysis 67
Results 68
Discussion 71

Chapter 4: 73

Illumination and Chromatic Colour Discrimination 73

Experiment 2.1 –Cone-Contrast Discrimination Test (CCDT) 77

Methods 77

Ethics and declaration 77

Participants 78

Design 78

Apparatus 78

Stimuli 79

Procedure 80

Thresholds 81

Experiment 2.2 – Illumination Discrimination 81

Methods 81

Design 81

Apparatus 82

Stimuli 82

Illumination Generation 83

Procedure 84

Obtaining Thresholds 85

Results 85

Discussion 87

Chapter 5: 89

The effects of Achromatic Surfaces on Colour Constancy..... 89

Experiment 3.1 93

Methods 93

Participants 93

Design 94

Stimuli 94

Procedure..... 96

Obtaining Thresholds..... 97

Results 97

Discussion..... 99

Chapter 6: 103

Colour constancy through illumination Discrimination, in real and simulated scenes.
..... 103

Experiment 4.1 105

Methods 105

Ethics 105

Participants 105

Design 106

Apparatus 106

Stimuli 108

Procedure..... 111

Thresholds 113

Results 113
Discussion 114
Experiment 4.2 118
Methods 118
Ethics 118
Participants 118
Design 118
Apparatus and Stimuli 119
Procedure 120
Results 121
Discussion 122
Chapter 7: 124
Immersive Colour Constancy 124
Experiment 5.1 126
Methods 126
Ethics 126
Participants 127
Design 127
Apparatus and Stimuli 127
Procedure 129
Thresholds 130
Results 130

Discussion 131

Chapter 8: 133

General Discussion 133

Are colour constancy mechanisms biased towards particular illuminations? 134

Can familiar objects cue colour constancy mechanisms towards particular illumination changes? 140

What information is needed for illumination discrimination? 143

What is the scope of these data? 144

References 149

Appendix 1. 159

Appendix 2. 161

Appendix 3. 162

Appendix 4. 163

A fellow named Brad, as we see,
Is as bright and as cool as can be;
He dispersed any dolour
With a lecture on colour;
And emerged with a new Ph.D.

- M Hammerton

Chapter 1:

An Introduction to Colour Constancy

The ability to sense a change in light irradiance and therefore the ability to detect changes in the environment from a distance is of such advantage to life that it is observed in species from the Cambrian explosion (Parker, 1998). Vision equipment has evolved independently between phyla, with anthropods, molluscs and vertebrates each possessing different, convergently evolved, eyes (Ogura, Ikeo, & Gojobori, 2004). Most seeing animals which possess a single class of photoreceptor are able to detect only changes in irradiance within one narrow band of the electromagnetic spectrum, while some diurnal species possess more than a single photosensitive pigment; these animals have colour vision, and are able to discriminate surfaces of equal surface radiance (Gegenfurtner & Kiper, 2003).

Humans most often possess trichromatic colour vision, having three photosensitive cone pigments (photopsins) sensitive to short– (S-cones), middle– (M-cones) and long– (L-cones) wavelength light, with peak sensitivities near 442, 543 and 570 nanometers (nms) (Stockman & Sharpe, 2000). These cones signal photon absorption by hyperpolarising, resulting in the increased firing activity of connected ganglion cells; however, once a photon has been absorbed by the cone opsin the wavelength information is lost, a principle known as univariance (Rushton, 1972). A ganglion cell with a high firing rate could therefore be signalling many photon catches of wavelengths with low affinity to the connecting cone's sensitivity function or fewer photons of wavelengths at that cone's peak sensitivity (Rushton, 1972)..

A consequence of univariance is that the perception of a patch of light in the void, composed of monochromatic yellow light at wavelength 580nm, can be matched with light composed of mixtures of monochromatic red (~680nm) and green (~545nm) light such that the L— and M— cone quantal catches are equal, a principle used in Rayleigh matching (Rayleigh, 1881; Rushton, 1972; Thomas & Mollon, 2004). It is therefore true that any two spectra that excite the cones equally will be perceived as the same colour, despite having differing spectral compositions, a concept defined as metamerism (Hunt, 1991; Wyszecki & Stiles, 1982). Furthermore, because the light reflected from surfaces in any scene can be characterised by the illumination spectral power distribution, $I(\lambda)$, multiplied by the surface reflectance function (SRF), $S(\lambda)$, it is possible that two surfaces can have the same surface colour, $E(\lambda)$. This is achieved when the combination of the surface reflectance function, the illumination spectral power distribution and the spectral sensitivity function of the photoreceptor classes, $R(\lambda)$, yields the same result, see Formula 1.1.

$$E(\lambda) = \int_{380}^{780} R(\lambda)I(\lambda)S(\lambda)\ d\lambda$$

Formula 1.1. The sensation (*E)* can be defined as the multiplication of the spectral content of the illumination (*I)*, a surface (*S)* and the photoreceptor sensitivities (*R)*.

However, two surface colours reflecting the light that would usually appear the same in the void can elicit the perception of two different colours, as seen in Figure 1.2. The king chess piece in the centre of each scene is reflecting the

same light, as the illumination spectrum and surface reflectance function of the central piece in each image were switched before rendering; that is, the signals from the two objects are metameric, yet we perceive the left king as bluish and the right king as yellowish. This means that colour perception is driven, at some level, by context within the image (Hurlbert, 1996).

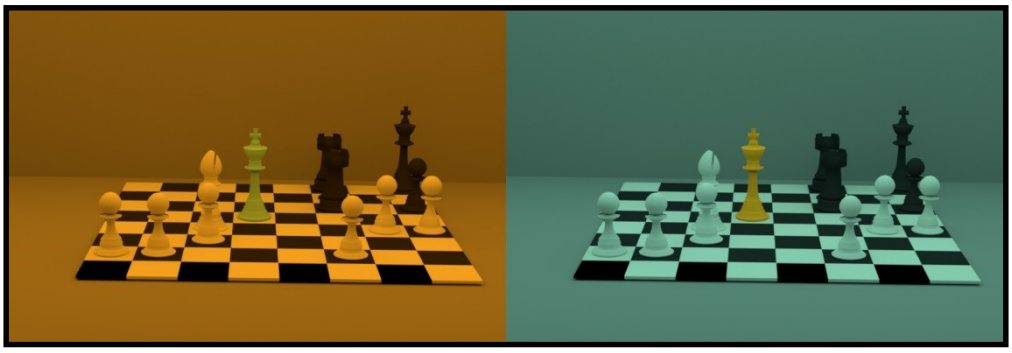

Figure 1.2. The king in the centre of each image are reflecting the same light, however one appears blueish (left) and the other yellowish (right).

While surfaces reflecting the same light can appear different, as in the above example, so can object surfaces appear the same colour when the light they reflect changes. Indeed, as the illumination on those surfaces changes, so does the light reflected; however, the perception of surface colours remain roughly stable, a phenomenon known as colour constancy (Foster, 2011; Gegenfurtner & Kiper, 2003; Hurlbert, 1999; Smithson, 2005). A demonstration can be seen in Figure 1.3; as the illumination changes from blue to yellow, many surfaces such as the bananas, appear to stay roughly stable. A cut-out marked A, shows that the average of the bananas in the left scene is actually green, yet they appear yellow. The cut out is copied to both scenes to show that the colour of the bananas in the left scene actually most resembles the pear in the right.

Moreover, the cut-out appears different depending on whether it is placed in the right or left scene, in the left hand scene it appears yellow in the right green.

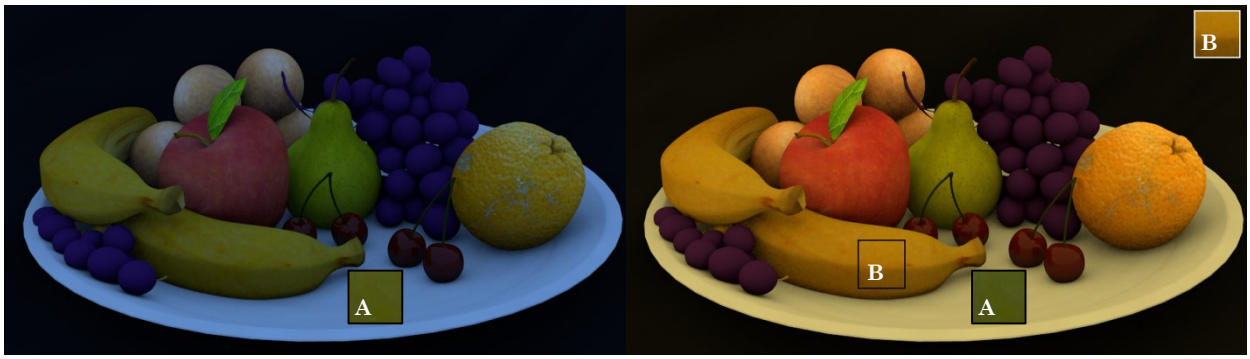

Figure 1.3. Bowl of fruit rendered using two illuminations, bluish to the left and yellowish to the right. The cut-outs marked (A) are the average colour of the left banana, under the blue illumination. The cut-out marked (B) has the same average colour as the orange in the same (right) scene, despite the banana appearing yellow.

Often, as can be seen in the example described above, colour constancy is not perfect. That is, the banana does not look exactly the same yellow, but one would still usually use the same colour name to describe it (Smithson, 2005). Some surfaces are completely different, such as the bowl in both scenes. A perfectly colour-constant observer would perceive both scenes as if surfaces were illuminated by an equal-energy white illumination, regardless of whether the scene is illuminated by bluish or yellowish light. How colour constancy is achieved, when the illumination information is not available to the visual system remains elusive (Foster, 2011; Hurlbert, 1997; Pearce, Crichton, Mackiewicz, Finlayson, & Hurlbert, 2014).

Understanding how colour constancy mechanisms operate, when they break down and why they are imperfect is vital to understanding why such mechanisms exist in humans, and other animals -- and is critical in quantifying our colour perception.

Quantifying Colour

Before colour appearance can be effectively studied, "colour" itself must be quantified. All colours can be represented with a three element array, as only three degrees of freedom are required to quantify trichromatic human colour vision (Grassman, 1853; as in Gegenfurtner & Kiper, 2003); this array, E, is called a tristimulus value and is generated by integrating the pointwise multiplication, with respect to wavelength, of known spectra over three independent functions, R; as in Equation 1.1. These tristimulus values represent how much light, using spectra comprised of those matching functions, is required to match that colour; those functions may be the human cone sensitivities, or any three other independent, primary functions (Stockman & Sharpe, 2006); Figure 1.4 shows the human cone fundamentals and CIE 1931 imaginary XYZ colour matching functions side by side (Stockman & Sharpe, 2000; Wyszecki & Stiles, 1982). Each differing array, E, is a unique chromaticity, and the geometric representation of all chromaticities together produces a colour space (Wyszecki & Stiles, 1982).

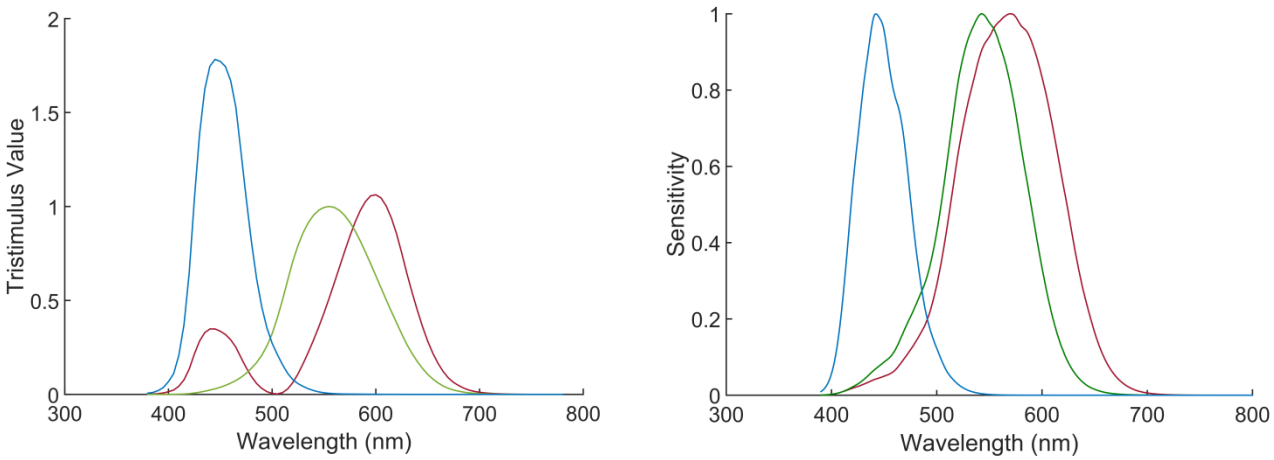

Figure 1.4. Left, the CIE 1931 imaginary XYZ colour matching functions (Wyszecki & Stiles, 1982). Right, the Stockman and Sharpe (2000), 2deg cone-sensitivity functions.

Various colour spaces exist, with differing units, defining different properties of the colour being described. However, all colours in one colour space can be represented within another, as they are produced by colour matching functions, and each coordinate represents some percept. E.g. The CIE 1931 Yxy colour space uses three colour matching functions, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ – that when integrated with a spectral power distribution on a wavelength by wavelength basis, result in X, Y and Z tristimulus values which are used to position that chromaticity geometrically as a function of those colour matching functions; it was developed by the *Commission Internationale de l'Éclairage* in 1931 (Hunt, 1957; Wyszecki & Stiles, 1982); a plot of the CIE 1931 colour space can be seen below in Figure 1.5; notice that the plot is two-dimensional, this is because chromaticities are represented without luminance (Wyszecki & Stiles, 1982).

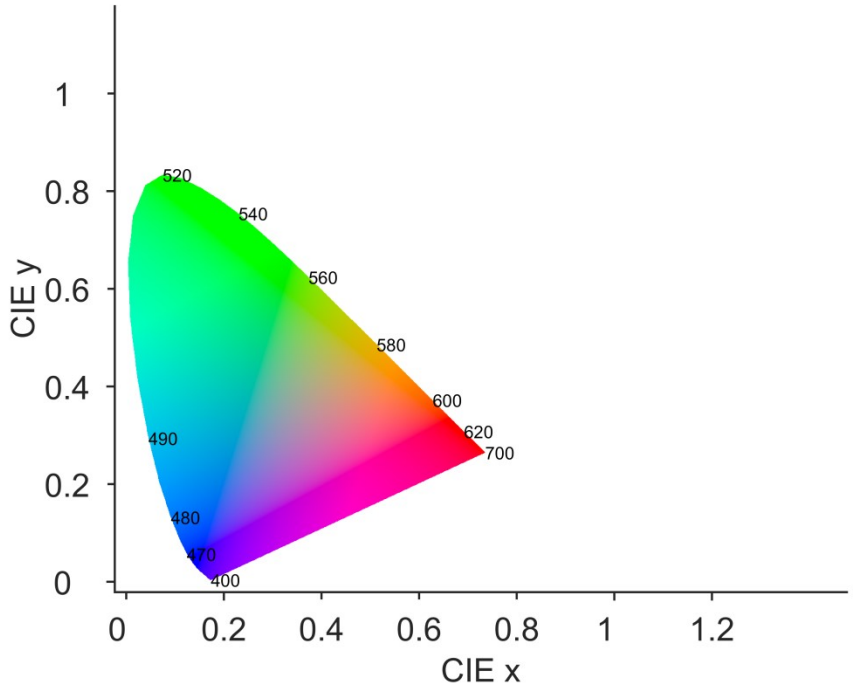

Figure 1.5. The CIE 1931 chromaticity diagram with simulated colour fill; numbers mark the coordinates of monochromatic light at respective wavelengths.

The CIE 1931 colour space is most common for describing colour, however it is not perceptually uniform; that is, a single unit step (a 0.01 step, for example) in one area of the colour space may correspond to a different number of discriminable colours than in another portion of the same colour space; this can be seen by the discrimination ellipses described by MacAdam (Wyszecki & Stiles, 1982). Therefore, to quantify the perceptual relationship between points in colour space, as a function of geometric distance, further transformations, such as those derived from the MacAdam ellipses, can be applied to produce more perceptually uniform colour spaces (Wyszecki & Stiles, 1982); an example is the Lu^*v^* colour space, which will be discussed later, and is documented extensively elsewhere (Schanda & International Commission on Illumination., 2007).

The Nature of the problem

Mechanisms that mediate colour constancy do not have information on the illumination directly, and to achieve constant surface colours the illumination must be discounted (Foster, 2011; Hurlbert, 1989; Smithson, 2005); the colour of the illumination must be derived from cues within the visual scene, and perhaps prior information about the possibilities of which illuminations can or should occur (Finlayson, Hubel, & Hordley, 1997). Effective models of colour constancy must therefore describe the equations necessary to transform the input data from the sensor (quantified by tristimulus values), to data as if the scene was sensed under an illumination of equal energy across the spectrum (computed equivalent tristimulus values). Efforts have focused on treating this as: 1) a purely computation problem; 2) by studying the physiology of nervous systems capable of colour constancy, and 3) through visual psychophysics (Foster, 2011; Gegenfurtner & Kiper, 2003; Hurlbert, 2003; Smithson, 2005).

The processes mediating human colour constancy begin with the computation in the retina; where a colour-opponent system contrasts cone inputs to create two chromatically contrasting axes (D'Zmura & Lennie, 1986; Hurvich & Jameson, 1957); S-cones are compared against the sum of the other two cone types (S – [L+M]), producing a +blue and –blue (blue-yellow) colour axis and L-cones and M-cones (L – M) are compared to produce a red-green axis. Physiological evidence shows S-cones connect via S─ON bipolar cells to small bistratified ganglion cells which also take input from L+M─OFF bipolar cells producing a chromatically, but not spatially opponent blue-yellow channel (Dacey & Lee, 1994; Dacey, 1996; Lee, 2014). The poor spatial resolution of the

blue-yellow system is supported by the scarcity of S-cones in the fovea (Curcio et al., 1991).

The red-green channel is both spatially and chromatically opponent at the central and para-central fovea, with a single cone class (L or M) connecting to the centre of each midget ganglion cells' receptive field by either an ON or OFF midget bipolar cell, with a mixture of L and M cone classes connecting to the surround receptive field via ON or OFF midget bipolar cells (inverse to that of the centre) (Dacey, 1996; Field et al., 2010).

The signals that leave the retina have undergone processing which is both spatially and temporally sensitive, and undergo further processing at the koniocellular and parvocellular layers of the lateral geniculate nucleus (LGN); while the wiring of the LGN is largely capable of only luminance contrast (Conway, 2009), cells have been isolated that are both spatially and chromatically sensitive (Conway, 2013; Lee, 2014); however, the role of, and a predictive model of, the LGN's role in colour vision remains elusive.

Colour constancy mechanisms appear to be observable in V4 as described in the Rhesus monkey by Zeki (1980), showing cells responding to surface colour not the composition of the spectra illuminating them (Foster, 2011); that is, cells whose response is consistent with our percept of surface colour, as described by psychophysical data. However, further lesion studies, in monkeys, have cast significant doubt on V4 being the centre of human colour vision, with monkeys with V4 lesions able to sort chromatically varied tiles (Heywood, Gadotti, &

Cowey, 1992); suggesting that further cortical sites are responsible for colour perception.

The great difficulty in isolating cells that perform colour processing is identifying cells that respond to our percept of colour, specifically colour in context (Hurlbert, 2003); that is, isolating cells that respond differently to the two patches marked 'A' (in Figure 1.3) depending on their surrounding context, despite them being identical in chromaticity. Moreover, to fully qualify colour constancy, the cellular network which brings about stable colour appearance must be characterised and therefore the cues which inform these processes must be isolated.

Various cues, such as the scene average chromaticity, and the brightest part of the scene have been proposed as properties that could be used by the retina and the cortex to cue colour constancy mechanisms (Hurlbert, 1989; Land, 1977). Computational models of colour constancy known as lightness algorithms propose models to transform the sensory input using these statistical properties to achieve colour constancy (Barnard, Cardei, & Funt, 2002; Hurlbert, 1989, 1998; Smithson, 2005).

The earliest model of colour constancy was proposed by von Kries and formulated by Ives (1912), and is often referred to as von Kries adaptation (in Hurlbert, 1998). The model proposes a linear transformation be applied to each of the cone-input channels independently, to attempt to discount the chromaticity of the illumination, whilst preserving the ratios of those cone excitations. Indeed, spatial cone-excitation ratios, the relative extent to which

different surfaces elicit responses from a particular photoreceptor class, have been shown to be largely invariant under natural illuminations; however, it appears that long adaptation times (> 1 min), or two-stage processes are required to achieve good constancy; and thus, cone-excitation ratios are not suitable for explaining the adaptation to very rapid illumination changes experienced ecologically, with skylight illumination changing continuously (Dannemiller, 1993). The preservation of these ratios has been offered as an explanation for relational colour constancy, the phenomenon that the relationship between surface colours can be used to identify if a change in surface colour is due to an illumination change or a change in surface reflectance function (Foster & Nascimento, 1994; Foster et al., 1997). An example can be seen in Figure 1.6: the left and middle image appear to be of the same scene shown under two different illuminations, bluish and yellowish illuminations respectively, as the ratios of the cone excitations have been preserved in the two images, although the absolute values have been shifted; the right hand image appears to be of a different scene, under an indeterminate illumination relative to the left hand image, as the cone ratios have been violated.

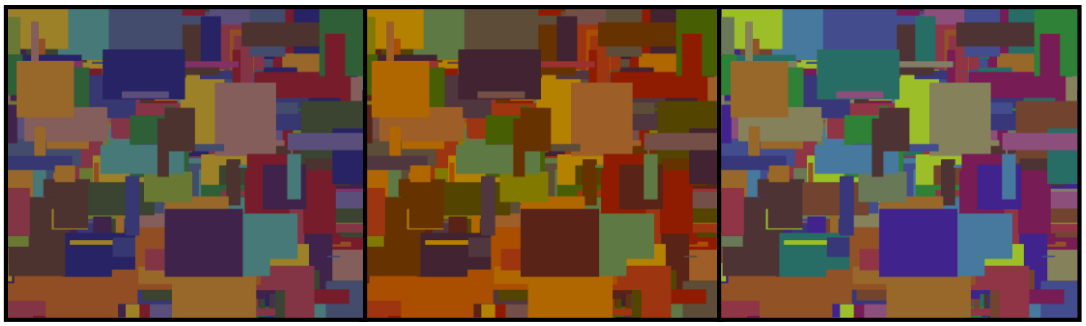

Figure 1.6. Three Mondrian scenes containing patches with colour checker chart reflectances. The left image and middle images have had the chromaticity of each patch modulated, by multiplying x by a constant factor (CIE Yxy 1931 space). In the right image, the chromaticity of each patch has been modulated by multiplying x or y by random amount.

The preservation of cone ratios alone cannot fully explain colour constancy. As can be seen in Figure 1.3, surfaces only remain constant if their cone excitation ratios remain the same relative to each other within the context of the scene, when a patch is devoid of context the patch appears as it would in the void (Patch A); therefore, spatially distinct patches need to be segmented first (Foster, 2011), suggesting more complex computations are involved than transforming the cone-excitations. Moreover, some surfaces within a scene may be metameric under one illumination, but not under another, which necessitate a violation of cone-excitation ratios (Smithson, 2005). Indeed, in Figure 1.3, the banana and orange in the right hand scene are roughly metameric; however one appears yellow and the other orange (see patches marked B).

Because simple von Kries scaling does not offer a solution to determining the illumination chromaticity spatially, further advances were proposed by Land (1977) in his Retinex algorithm; so named to highlight that both retinal and cortical processes contribute to colour constancy. There are two main versions of the Retinex algorithm, based on two different assumptions. The first is the brightest-is-white (max flux) hypothesis, which assumes that that brightest part of the scene should be reflecting the composition of the illumination most accurately and the input sensors should be scaled relative to that point. The second is that the spatial average of the scene could be assumed to be grey, and deviations from the average can be used to estimate the illumination (Land, 1986), known as the grey-world hypothesis.

Retinex has been shown to work well for some scenes (Hurlbert, 1989). However, the assumptions can be violated easily in some natural images as many have demonstrated (Jobson, Rahman, & Woodell, 1997) – scenes are rarely grey on average under neutral illuminations, and natural illuminations are rarely uniform enough for linear scaling to be effective without compression of the dynamic range within the image, irrespective of environment.

Nascimento and Foster (2000) describe our ability to discriminate the relationship between surfaces in isoluminant images, where there is no brightest point; observers were able to detect illumination changes over surface changes, where the only cue was spatial cone-excitation ratio modulation, suggesting that 'brightest is white' does not describe human illumination change discrimination ability. Moreover, these lightness algorithms do not allow for prior information

to be considered, or a constrained set of illumination possibilities (Finlayson et al., 1997). Nevertheless, lightness algorithms make predictions about the nature of human colour constancy that are testable experimentally.

Experimental Approaches to Colour Constancy.

Colour constancy is typically measured experimentally by equating the perception of some colour in a scene to a chromaticity within a colour space. That is, the measurement describes the perception of the light reflected from a surface as if that light were perceived in the void, devoid of any other context. For example, we could ask an observer to make a square patch of colour on a computer screen appear the same as the banana or orange does in the bowl of either the left or right images in Figure 1.3. Once the observer is satisfied with the match, we can plot the chromaticity of those matches in a colour space along with the actual chromaticity of the light reflected from the objects. This technique is called matching by adjustment (Smithson, 2005; Foster, 2011).

Arend, Reeves, Schirillo and Goldstein (1991) asked participants to adjust a patch of colour on one monitor (the test patch) to match another (the reference patch) in the same position on another monitor; this method is called simultaneous colour matching. Two scenes were used: one where both the test and reference patches were positioned in a scene containing a uniform coloured background, and another where the patches were imbedded in a variegated scene of many other coloured patches, commonly referred to as a Mondrian (see Figure 4). Each reference scene was under one of three daylight illuminations, yellowish (4000K), bluish (10000K) and neutral (D65, 6500K), and the test scene

was held under D65. After adapting to a D65 field for 3 minutes, observers were asked *either* to match the test patch to the reference in hue, saturation and luminance, *or* to match the patch to "look as if it were cut from the same piece of paper," Arend et al., (1991). They then transformed each match's RGB coordinates into the CIE 1976 u'v' colour space, which is considered perceptually uniform for an observer adapted to D65 illumination (Hunt & Pointer, 2011). The perceptual distance, $\Delta E_{u^*v^*}$, of each match from the coordinates of that patch under the neutral illumination was then computed, as can be seen in Equation 1.5; this is the perceptual shift. The physical distance is defined as the distance from the patch's physical chromaticity coordinates under the new illumination and under the neutral illumination. The ratio between the perceptual shift and the physical shift of the test patch's colour was then calculated (see Equation 1.6), to yield a Constancy Index as a yardstick for measuring the level of colour constancy (Arend et al., 1991; Brainard, 1998). Higher constancy was reported for matches when the observer was told to match the test patch as if it were the same piece of paper (mean CI: 0.52), than when they were asked to match the hue, saturation and luminance in the reference scene (mean CI: 0.2). The authors reported no differences between the uniform and Mondrian scenes.

$$\|\Delta E_{u^*v^*}\| = \sqrt{\left(E_{1_u} - E_{2_u}\right)^2 + \left(E_{1_v} - E_{2_v}\right)^2}$$

Equation 1.5. Formula for calculating the Euclidean distance, $\Delta E_{u^*v^*}$, for the two chromaticities $E_{1\,..2}$; the Euclidean distance is equivalent to a number of perceptual steps, $\Delta E_{u^*v^*}$, in discriminability or the perceptual distance, between E_1 and E_2 (Brainard, 1999; Arend et al, 1991).

$$CI = 1 - P(\Delta E_{u'v'}) / E(\Delta E_{u'v'})$$

Equation 1.6. The ratio between the perceptual shift in the colour of a patch, E, and the physical shift, P, when the illumination changes from neutral yields the Constancy Index (CI), (Brainard 1999; Arend et al, 1999; Smithson, 2005; Foster, 2011).

Arend et al., (1999) and others (Arend & Reeves, 1986; Reeves, Amano, & Foster, 2008) suggest that there must be some mechanism for retrieving the surface reflectance of the patch when it is under an arbitrary illumination, as constancy indices were higher in the condition where observer where asked to make a 'paper' match. Moreover, there appeared ability for observers to make faithful matches to the physical change; meaning that observers had some control over colour constancy; however, constancy indices were rarely close to 1 (perfect colour constancy, or no perceptual shift) or exactly 0 (no colour constancy). This demonstrates that higher-level cognitive concepts, such as the constancy of an object being illuminated by changing illuminations, could contribute to levels of colour constancy; these will be discussed later.

Brainard and Wandell (1992) describe an asymmetric colour matching task, in which the observer was exposed to an array of coloured patches (a 5x5 array of colours from a pool of 226 reflectances), against a uniform background, under a large number of illuminations (a training set). Observers then saw a test scene, under a test illumination. Observers were then asked to make a test patch match one of the patches on the same monitor once the scene was not in view.

This removes mixed states of adaptation experienced when observers were simultaneously matching, but does introduce a memory component (Brainard & Wandell, 1992). They found that standard perceptual transforms (Linear, Diagonal and Affine) on the tristimulus values of observer matches could predict settings within an approximate room mean square error of ~7 perceptual steps.

These chromatic matching tasks assess the appearance of a small number of chromaticity coordinates within the context of the scene, under a set of given illuminations; however, the focus is on the appearance of the chromatic patch and not the general state of adaptation of the visual system to the illumination, which is not measured directly. Instead an indication of constancy for each surface is assessed, as described above.

To attempt to measure the adaptation point of the observer directly, Brainard (1998) describes a technique called achromatic adjustment. In this task observers adjusted a patch until it appeared achromatic, containing no blue, yellow, red or green, on a continuum between black and white (Brainard, 1998; Chichilnisky & Wandell, 1995). The rationale was that a patch that appears white to an observer, who is perfectly adapted to the scene illumination, should match the illumination chromaticity, because a white patch should perfectly reflect the illumination spectrum. The observer's achromatic setting would thus reveal their state of adaptation. Furthermore, as observer's matches rarely indicate perfect colour constancy, the match can indicate the 'equivalent illumination' , that the observer's visual system assumes illuminates the scene; this is modelled by generating surface and illumination reflectance functions

that would result in the matches provided by the observer (with certain constraints) (Brainard & Maloney, 2011).

Kraft and Brainard (1999) designed four real scenes to test the contributions of the local surround, spatial mean (grey-world) and max flux (bright-is-white) to colour constancy. Two scenes were used in the local surround condition, both scenes containing a viewing box lined with card, some grey paper shapes, tube wrapped in tin foil and a colour checker chart. The background light was equated across both scenes by using grey card lining and neutral illumination in one, and a blue card lining and a reddish illumination in the other. This meant that cues from local surround were silenced. In the spatial mean condition the same two scenes were used but the spatial average of those scenes were equated such that cues to the illumination from the spatial average were silenced. In the max-flux condition, the other objects were removed from the viewing box and dark grey and yellow card lining were used under the neutral and yellow illuminations, respectively. This time however, a coloured card border surrounded the test patch which was held constant and was the brightest part of the scene, thus silencing cues from the brightest part of the scene. The test patch was illuminated by a projector, programmed to move in steps in the La*b* perceptually uniform colour space. Each observer adjusted the test patch in each condition until it appeared achromatic.

Kraft and Brainard (1999) calculated constancy indices for each condition, where an index of 1 indicated perfect colour constancy with an achromatic match that matched the illumination chromaticity, and 0 colour constancy where a

match was white under a neutral illumination. The mean constancy index for the local surround condition was 0.53 and for the spatial mean and maximum flux of 0.33. A control experiment where no cues were silenced yielded mean constancy indices of 0.83. These results indicate that the assumptions underlying classical computational theories, grey-world or max-flux, cannot fully account for colour constancy under natural viewing conditions (Kraft & Brainard, 1999); if they could, constancy indices would be 0 or near to 0 when these cues were silenced. On the other hand, the information silenced does contribute to mechanisms of colour constancy, as can be seen by the reduction of constancy indices when these cues were silenced.

Both chromatic and achromatic matching tasks are subjective measurements of colour constancy, measuring the appearance of surface colours or the adaptation point of the observer, under particular illumination changes. While these matching tasks give accurate measures of constancy between illumination changes for each observer, it becomes difficult to compare adjustments between observers (Foster, 2011). Moreover, the adjustment paradigm can be biased towards the localised contrast of the adjustment patch rather than the entire scene (Foster, 2011), making it difficult to compare constancy indices between experiments, especially as when a patch is being adjusted the surround contrast is continuously being modified (Foster, 2011).

Other measures of colour constancy have been developed that are objective, in that the aim is not quantifying the colour appearance of surfaces, but determining the ability of the observer to perform a forced-choice based on the

information within a scene. Craven and Foster (1992) define operational colour constancy as the ability to discriminate between a change in the illumination on the surfaces within a scene and a change in the surface reflectance functions of those surfaces. Observers saw a Mondrian scene where all surfaces changed in the x direction of the CIE 1931 Yxy space by a fixed amount (a global illumination change), or select surfaces changed by some amount (surface reflectance change). Participants were able to discriminate between illumination changes and surface reflectance changes above chance (d' > 0) for changes in the Mondrian patch chromaticities of Δx = (-0.06, 0.05 … 0.06); and subsequently under various controlled experiments. Relational colour constancy (Foster & Nascimento, 1994), through preservation of cone excitation ratios, has been suggested as closely linked perceptual phenomenon (see Foster, 2011 for review). The model of operational colour constancy states that if the observer attributes a change in the scene to the illumination, then the observer's visual system must perceive the surfaces as roughly constant. Importantly, this judgment is different to a judgment of colour appearance. The observer that attributes the change to an illumination change is exhibiting some level of constancy; however, the actual appearance of each patch is not measured.

It has been shown that the level of information within a scene, or the level of chromatic complexity, known as 'articulation' , can effect operational colour constancy (Maloney & Schirillo, 2002). Zaidi & Smithson (2004) describe a scene of a single patch under an illumination. If there is a change in that patch's colour, that patch could have changed colour or the illumination could have changed colour, and it would be impossible for the visual system to determine.

For example, in Figure 1.7, the two patches to the left could be both neutral under different illuminations or coloured under a single neutral illumination. It becomes clear that they are both neutral when billiard balls are added to the scene, in the right pane. Notice, the presence of the billiard balls makes the tiles appear whiter, even though the illumination on the scenes remains unchanged; that is, the tile is remaining colour constant even though the tile has never been seen under a neutral illumination.

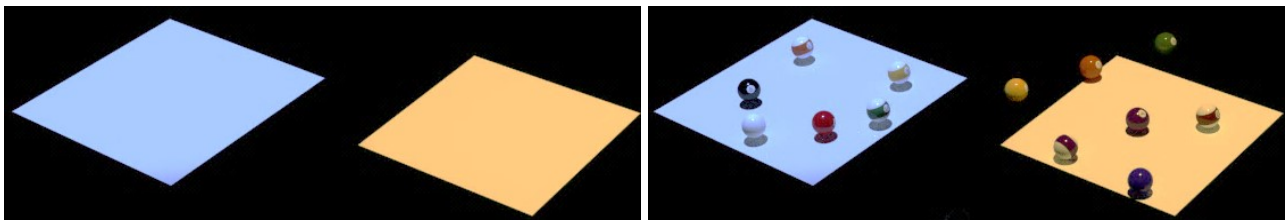

Figure 1.7. An example of colour constancy as a function of scene complexity: in the left image it is impossible to tell whether the two squares are blue and yellow, under a white light, or white tiles under a blue and yellow light. The right image indicates that the tiles are probably white, under two differing lights; as a result, the floor tiles appear lighter.

Linnell and Foster (2002) describe a forced-choice paradigm in which observers decided whether a change in a scene was due to a change in illumination or a change in surface reflectance. Mondrian scenes with varying number of patch chromaticities (between 1 and 49 and 50000) were used in each trial for the respective changes. As the number of patches increased in the scene, the accuracy of observers in detecting the correct change also increased. The frequency of low-level colour information, and the systematic way that

colour information changes, can cue the visual system to an illumination change (Foster, 2011) and subsequently the illumination colour (Amano, Foster, & Nascimento, 2006; Foster, Amano, & Nascimento, 2006). The evidence from scene articulation and operational colour constancy sets a lower limit for colour constancy, that a minimum of two co-varying surfaces are required to resolve a change in illumination (Smithson, 2005).

Aside from frequency of surfaces, there is evidence that a single uniform surface (in addition to a target surface) can cue constancy mechanisms to an illumination change (Hansen, Walter, & Gegenfurtner, 2007). Olkkonen, Hansen & Gegenfurtner (2009), had participants identify simulated disks of uniform chromaticity on a computer monitor under changing illuminations, by signalling it as belonging to one of eight different colour categories . This colour naming technique measures colour constancy using the frequency a colour name is given to the disk (Foster, 2011; Jameson, 1983; Olkkonen et al., 2009). The frequency that a disk, with fixed surface reflectance (a Munsell colour), was identified as the hue it appeared under a neutral illumination, was used to rate colour constancy, under changing illuminations; a constancy index of 0 was obtained when a patch was never given its correct colour name and an index of 1 when it was always given its correct name, regardless of the illumination change. The mean constancy index for observers was 0.8 when a grey background was visible, which fell to 0.65 in reduced cue conditions, when the background was silenced. This suggests that the mere presence of a neutral surface within a scene could cue colour constancy mechanisms (Foster, 2011).

Kraft and Brainard (1999), as previously discussed, reported constancy indices as high as Olkkonen, Hansen and Gegenfurtner (2009). As both experimental setups used neutral surfaces, it remains unclear whether colour constancy mechanisms may work better for real scenes than for simulated scenes (Foster, 2011). Foster (2011) compared the constancy indices of several aforementioned experiments using real scenes and simulated scenes and concluded that there appears no evidence to suggest that constancy indices are higher for real than simulated scenes, but that tasks differ so much between the experiments that it is hard to compare the indices directly. Therefore, the difference between colour constancy for real and simulated scene has not been systematically tested.

The fundamental literature here has focussed on developing measures of colour constancy. They have focussed mostly on controlling the scene statistics (Kraft & Brainard, 1999; Linnell and Foster, 2002), rather than investigating which surfaces are optimal for colour constancy (Foster, 2011). These experiments usually use a small number of illuminations, chosen arbitrarily, and surfaces that are mostly flat and uniform in colour (Arend et al., 1991; Kraft & Brainard, 1999; Foster, 2011; Smithson, 2005).

These studies have not focussed on the behaviour of the visual system in response to the cues available in the ecological environment under which these mechanisms of colour constancy have evolved, and have not determined whether these mechanisms are optimal under those conditions.

Colour Constancy and the Ecological Hypothesis.

There is substantial evidence that the visual system, starting at the retina through to the cortex, is optimised for natural scenes (Parraga, Troscianko, & Tolhurst, 2005; Regan et al., 2001; Sumner & Mollon, 2000). Primate colour vision appears to be optimised for discriminating fruits from foliage (Regan et al., 2001). Moreover, the illuminations under which primates have evolved, namely daylight, are regular and follow typical variations, as defined by the Planckian locus (Wyszecki & Stiles, 1982).

Predominantly daylight illuminations vary along the blue-yellow axis, and are predominantly blue (Hernandez-Andres, Romero, Nieves, & Lee Jr., 2001), because the probability of short wavelength light being scattered by the atmosphere is higher than that of long wavelength light (Wyszecki & Stiles, 1982). As a consequence of this, the principal component of variation in natural images is ubiquitously along the blue-yellow colour axis (Webster, Mizokami, & Webster, 2007); even for dense forest scenes with little skylight (Sumner & Mollon, 2000). Judd (1940) remarks that chromatic adaptation is almost complete under daylight illuminations, however the means to test this experimentally was not available.

Delahunt and Brainard (2004) had observers make achromatic adjustments of a test patch while viewing simulated scenes under eight illuminations. Four were extreme blue and yellow daylight illumination changes (60 and 30 $\Delta E_{u^*v^*}$ away from D65 for each colour direction), and four were red and green illumination changes of equal perceptual distance. Constancy indices for these settings were not significantly different from each other, with indices ranging

from 0.67 to 0.81; indicating that, at least when a single illumination is present for a prolonged adaptation time, colour constancy mechanisms appear equally competent for natural and novel illumination changes.

Conversely, when observers are asked to make a patch appear white in the absence of an illumination, their matches vary along the blue-yellow colour axis, along the Planckian locus (Bosten & MacLeod, 2012), suggesting that observers' internal representation of white is biased towards the appearance of white surfaces under daylight illuminations, and that the upper limit of colour constancy may be achieved under daylight variations.

Xiao, Hurst, MacIntyre and Brainard (2012) used an interleaved, adaptive staircase procedure and asked observers to make achromatic settings of a central object that was under one of two extreme daylight illuminations (~60 $\Delta E_{u^*v^*}$ from D65, blue and yellow). Observers indicated if the central object (a glossy billiard ball, matte ball or matte disk), within a simulated checkerboard viewing box was chromatic, via a forced choice of either: redder, bluer, yellower or greener than white. The staircase procedure reduced the colour contribution of that colour axis to the object depending on the observer's selection; for example, the observer's indicating an object was redder than white made the surface greener by an amount determined by the staircase. Once the observer was alternating between colour terms the staircase stopped; that is, they had reached their achromatic point because the object appeared devoid of colour. They found no significant difference between constancy indices for the two illuminations; however, constancy indices were significantly different for the objects, but this was dependent on which cues were controlled in the scene. When the

background contrast (local surround, as in Kraft and Brainard, 1999) was left unconstrained, constancy indices were higher for matte disk than spheres, whereas indices were worse for the matt disks when the background contrast was silenced. This demonstrates that particular object properties in conjunction with scene cues can inform colour constancy mechanisms.

It is unclear whether object properties, or higher-level cognitive processes can cue colour constancy, such as the expectation that an observed banana should be yellow. Granzier and Gegenfurtner (2012) asked observers to match the colour of the illumination on a scene with a Munsell colour, using real scenes. There were three scenes: one containing uniform coloured papers, and two containing fruits (including a banana), with other objects either typically coloured (the congruent cue condition), or atypically coloured (an incongruent cue condition). Constancy indices were ~0.39 for the uniform papers and the incongruent condition, which rose to ~0.47 in the congruent cue condition. However, Kanematsu and Brainard (2013), using successive colour matching, found that matches to an image of a banana were not significantly different than matches to a uniform patch that matched that banana's chromaticity.

These studies mainly focus on surface colours, under a small set of illuminations. Furthermore, they use a small sample of objects in mostly simulated scenes. It remains unclear whether colour constancy is optimised for the illuminations under which we have evolved. It also remains unclear whether colour constancy operates better for highly complex, natural and familiar objects that we usually encounter, compared systematically with equally complex, chromatically matched novel objects.

Objectives

The objectives of the presented work were (1) to understand whether colour constancy mechanisms are biased towards natural illuminations, and in particular, daylight; (2) to measure the effects of familiar objects on colour constancy mechanisms, compared with chromatically matched novel objects; (3) to determine whether colour constancy mechanisms operate better for real rather than computer simulated scenes; and (4) to observe colour constancy mechanisms under a large array of daylight illuminations, and physically similar novel illuminations.

The experimental hypothesis was that colour constancy mechanisms would operate better under daylight illuminations, and would also operate better when familiar objects were visible in a scene. A secondary hypothesis was that the rich cues in natural scenes would elicit better colour constancy than chromatically matched simulated scenes.

Instead of investigating colour constancy by measuring the perception of surface colours, here colour constancy is measured via the perception of the illumination colour, through illumination matching; a novel method developed here and documented in the coming chapters. The principle is simple: if an observer cannot see a change in the scene when the illumination changes, the surfaces in the visual scene which are illuminated by that illumination must appear constant in appearance.

The predictions of this principle were that detection of an illumination change would be poorer for daylight illuminations than other novel, broadband illuminations due to their abundance in natural scenes; thus, better colour

constancy as the illumination change was not detected; furthermore, that the presence of complex, familiar objects in the scene would make a change in illuminations harder to detect, regardless of the illumination change. Finally, that higher colour constancy would be observed for real scenes illuminated by spectrally-tuneable LED light sources, than simulated scenes, under simulated illuminations on a computer monitor; due to the multitude of cues in real scenes as compared to simulated, matte patches. These predictions and the methodologies used to test them will be discussed in much greater detail in the coming chapters.

Chapter 2:
Materials and Calibration Methods

Setup Overview

Tuneable LED light sources are the main component of the experimental setup. Unlike most colour constancy experiments where the surface colours are manipulated by the observer, the scene observed was held fixed within each experiment, and the illumination on that scene was varied. Observers viewed these scenes by looking through a porthole in the front of a viewing box, and performed an illumination matching task; the illumination on the observed scene was changed by programming the LED light sources with the spectra to be presented in real-time (Finlayson, Mackiewicz, Hurlbert, Pearce, & Crichton, 2014; Mackiewicz, Crichton, et al., 2012). Light was mixed from multiple light sources by an integrating sphere, 1m in diameter, and then reflected down into the viewing box, with almost perfect uniformity. A schematic diagram of the experimental setup can be seen in Figure 2.1; details of the task will be explored extensively in Experiment 1.1 (see Chapter 3).

The experimental stimuli, consisting of the illumination spectra and the scene composition (the contents of the viewing box), were quantified and systematically varied between experiments. These elements of the apparatus and the methods of preparing and calibrating them are discussed in detail in the coming sections.

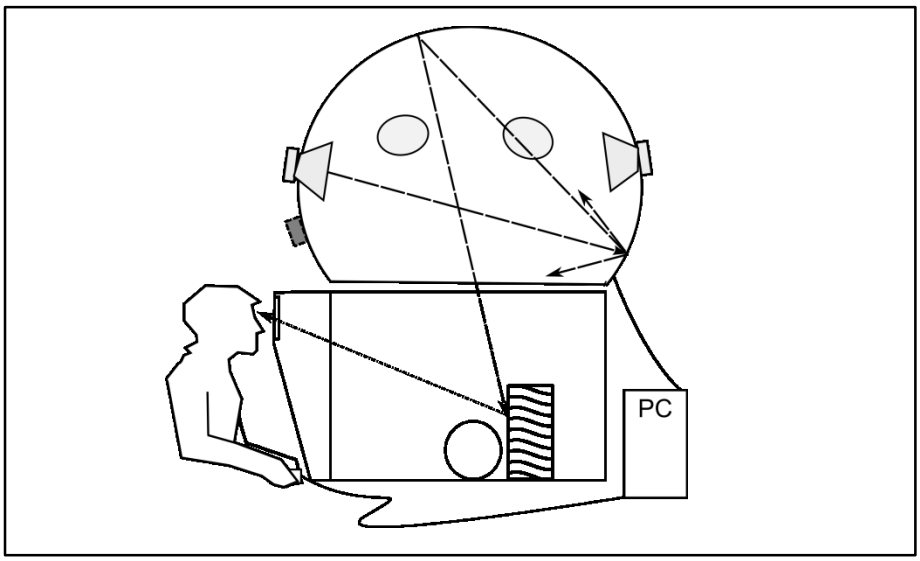

Figure 2.1. Basic experimental setup; a viewing box with a porthole allows the observer to see the box contents and supply feedback to a PC via a game controller. The PC controls the illuminations in the viewing box. Long-dashed lines show illumination from the integrating sphere, and short-dashed line shows the reflected light from scene contents. (Not to scale).

Viewing box

The viewing box had dimensions: 71cm (width) x 77cm (depth) x 47cm (height). A viewing aperture was in the middle of the box, 1cm from the top edge with dimensions: 7.5cm (height) x 14.5cm (width). The box was painted grey (x = 0.299, y = 0.324), and could be lined with paper as can be seen in Figure 2.2; in experiments where paper was used, each wall was lined. The top of the box was open to allow the light from the integrating sphere to uniformly illuminate the scene; an unpainted rim is visible (see Figure 2.2) which was covered with black cloth to create a light-tight seal between the integrating sphere and the box.

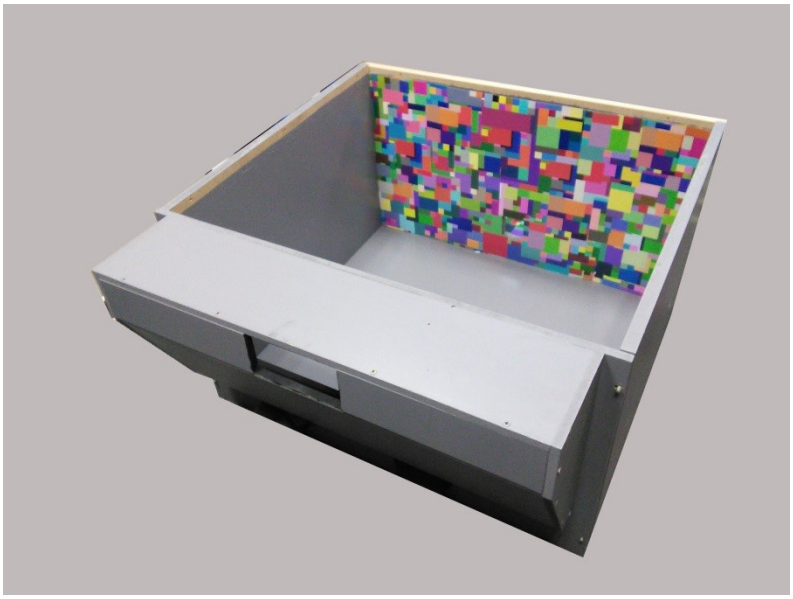

Figure 2.2. Photograph of grey viewing box; the back wall is lined with Mondrian paper. The top is open to allow illumination from above.

Tuneable LED Light Sources

The quantity of photons emitted from a light-emitting diode (LED) can be regulated by controlling the drive current to that LED (Van De Ven, Chan, & Wah, 2014), or by pulse-width modulation (PWM) to regulate the 'on' state (photon release) with respect to time. High-luminance LEDs have enabled the availability of computer monitors that use varying mixtures of the 3 primary LEDs (Red, Green and Blue) to produce a wide gamut of chromaticities, and general purpose light sources (Sheats et al., 1996). Whitish illuminations created with three narrow-band primaries (such as those created by commercially available RGB LED lamps) have peak intensities at those primaries. These produce a smaller gamut of chromaticities from surfaces within a scene than more broadband illuminations due to the light available at each wavelength; therefore, these solid state lightsources (SSL) typically also

contain a broadband yellow-phosphor LED which increases the colour-rendering index (CRI) of the lightsource, as a high luminance, green LED remains unavailable to fill the wavelength gap (Lin, 2010; Sheats et al., 1996).

The broadband channel of these SSL light sources limits the use of the LED lamps for specific colour vision research without the use of filters. Mackiewicz et al. (2012) present a tuneable 'Illuminator' for vision research, encompassing 6 Gamma-Scientific RS5B-light sources, that projects light within an integrating sphere, in turn reflecting diffuse illumination into a typical viewing box. Each lamp within the illuminator contains 9 different LEDs, 8 narrowband and 1 broadband (yellow-phosphors, see Figure 2.4 for basis functions). Each channels' drive current is tuneable and can be controlled, in real time, independently at 16-bit resolution, allowing almost any spectral composition to be specified. The following sections will discuss how to calibrate multi-channel LED lightsources generally, focussing on the two different LED illumination technologies used in subsequent experiments, the RS5B drive-current modulated light-sources and the IREC prototype PWM luminaires. Furthermore, methods for fitting spectra for colorimetric accuracy and finding spectra with specific tristimulus values and other spectral properties will be discussed.

Calibrating LED light sources.

The quantity of photon emissions increases, with respect to time, as the drive current to an LED increases. Given a number of LEDs with known spectral emission functions, and independent control of the drive current to each, linear combinations of those functions can be used to produce a desired spectrum (Finlayson, Mackiewicz, Hurlbert, Pearce & Crichton, 2014; Mackiewicz et al, 2012). Thus, once the relationship between the drive current and the light reflected from a perfect reflector to the observer is known, the tristimulus values of light at a point can be calculated. As the drive current increases, so does the temperature of the gap junction within the LED; thus, the emission SPD and peak wavelength changes; which have been effectively modelled using the Maxwell-Boltzmann distribution (Baumgartner, Vaskuri, Kärhä, & Ikonen, 2014; Kärhä, Vaskuri, Baumgartner, Andor, & Ikonen, 2013). Due to there being a nearly endless set of states for a tuneable LED system, and because the time for thermal equilibrium changes as a function of drive current, known basis functions for a set of approximate thermal states are required. A useful set of basis functions are those which accurately predict the spectral output of the LEDs, after some uptime, for an arbitrary illumination; this criterion was the standard for assessing a successful characterisation of the LED system. This accuracy can be assessed spectrally or colorimetrically; the accuracy of each calibration is described for each experiment, however an arbitrary tolerance of $2\Delta E_{u^*v^*}$, from the desired chromaticity of the illumination to be presented, was accepted as an accurate illumination match.

Methods

Apparatus

A spectrally tuneable illuminator was used, consisting of 6 LED (Gamma Scientific RS5B) light sources, each with a bank of 10 programmable LED channels (8 x narrow band, 2 x broadband), which projected into an integrating sphere, 1 meter in diameter, the interior of which was coated with Barium Sulphate; the sphere had a porthole for taking measurements that was situated at ~45° from the exiting aperture. A black cloth was used to block the aperture on the integrating sphere. A PR650 spectroradiometer on a tripod stand was used to take measurements from inside the integrating sphere. A Windows 7 powered workstation was used to control the PR650 and illuminator through the MATLAB software package, using custom functions (Appendix 1). Data cables were used to connect to the PR650 and the illuminator.

Subsequent calibrations were performed using two sets of 3 IREC prototype luminaires (Mark I and Mark II). Each Mark I luminaire contained 13 different programmable primaries (9 x narrow band, 4 x broadband); Mark II luminaires contained 8 different primaries. These luminaires were mounted in the ceiling of a room painted with white. A Minolta CS-2000 spectroradiometer was used to measure the radiance from a Formazin standard by Gamma Scientific (~.5m x ~.25m w/h).

Design

A repeated measures design was used, where each LED channel (LED basis functions 1-10/1-8/13 for IREC luminaires) was measured at intensities 1%, and

10% - 100% in steps of 10%. The dependent variable was measured radiance ($W \cdot m^{-2} \cdot s^{-1} \cdot nm^{-1}$) in the integrating sphere, which was measured from 380nm to 730nm at 4nm resolution using the PR650 spectroradiometer and 1nm resolution using the CS-2000.

Procedure

The integrating sphere's aperture was blocked using a piece of black felt cloth. A small calibration hole at the side of the illuminator was opened (size 2.5cm diameter), and the PR650 was positioned such that the light from the aperture was focussed into the lens of the spectroradiometer (see Figure 2.3). All lights were turned off in both the calibration room and the illuminator. A measurement was taken of the background visible radiation (black reading) in the dome. The control software then set each channel to the desired intensity (1%, 10%-100%), in isolation, and then triggered a command to the PR650 to take a reading. The data from each reading was returned to the software and stored (or control software logic see Appendix 1.).

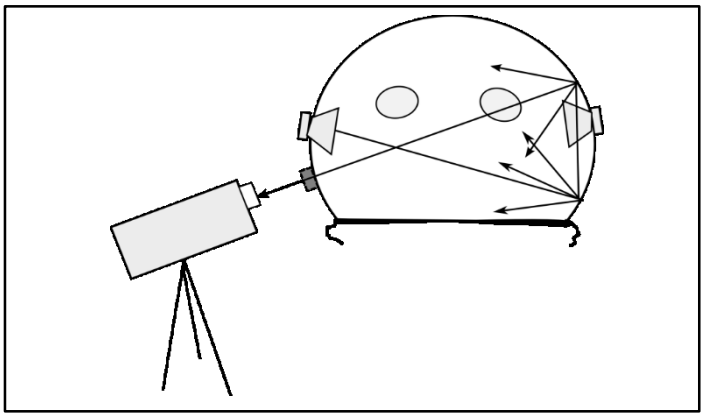

Figure 2.3. Diagram of spectroradiometer positioned near the porthole of the integrating sphere; with lamps radiation – leaving after successive reflections on the barium white lining (Not to scale).

The procedure using the IREC luminaries differed only in that the calibration tile was placed at the far wall of the room at 1m height with the spectroradiometer measuring through a porthole as before, cut in the near wall ~2m away from the reflectance standard. Due to the IREC luminaires using pulse-width modulation, only 1 reading per LED (power on full) was necessary, as current remained constant.

Results

The spectroradiometer reported that not enough light was available for the dark reading; under such circumstances the PR650 return the last successful reading, and therefore this was used to determine if a read error had occurred. No compensation was performed on the channel readings. The maximum intensities of each channel, scaled relative to that of the maximum of LED channel 2, can be seen in Figure 4.

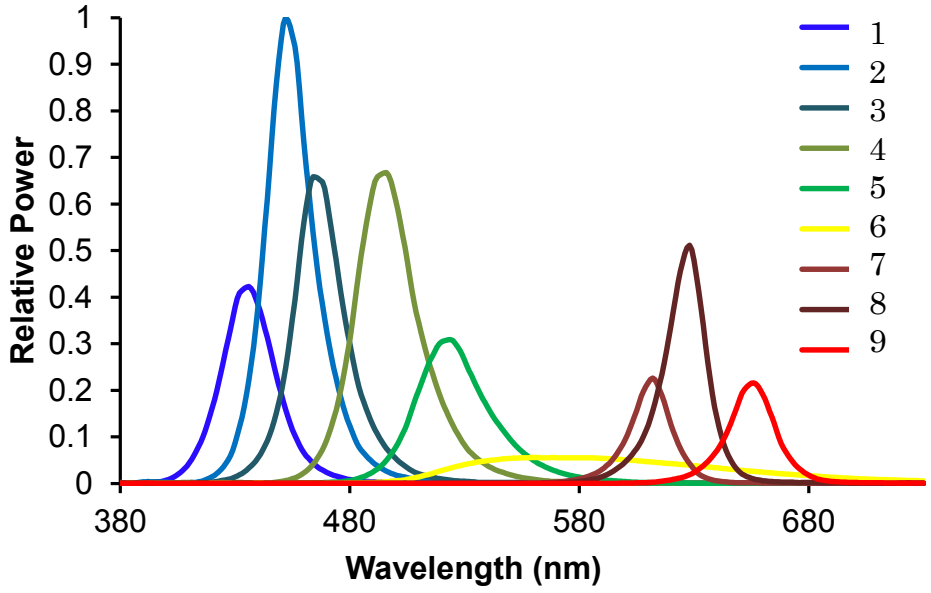

Figure 2.4. Each of 9 separate LED channels at maximum intensity scaled to that of channel 2. See legend for Channel numbers.

The peak wavelength (in nanometres) of each LED can be seen in Table 2.5. The mean maximal peak shift across all LEDs was 4.3 nm.

Table 2.5. Peak wavelength of each unique channel and maximal shift for RS5b luminaires.

LED	Max Intensity (nm)	Mean (nm)	Maximal Shift (nm)
1	436	436	0
2	452	455	4
3	464	467	4
4	496	498	4
5	524	528	8
6	560	564	13
7	612	612	0
8	628	628	2
9	656	655	4

Only LED channels 1 and 7 experienced no peak wavelength shifts as the drive current increased, with the shift for the yellow phosphor being most pronounced; its output is plotted below in Figure 2.6. These shifts are non-linear.

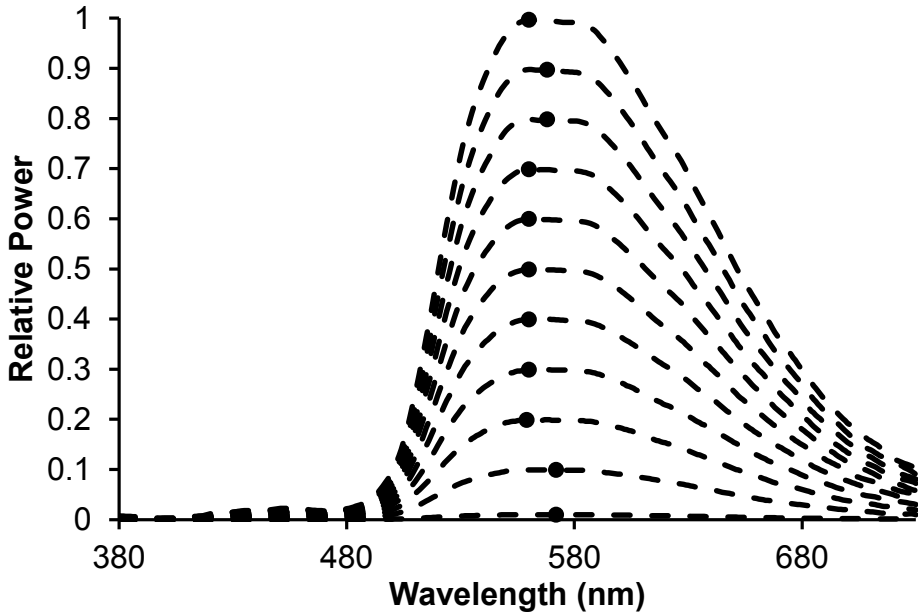

Figure 2.6. Relative spectral emission of LED Channel 6, scaled to power at maximal drive current. Black markers indicate spectral peak for each intensity band respectively.

For the IREC luminaries (Mark I & II), no peak shift was found as a function of current (which remained constant); however, temperature, which was dependent on the number of LED channels currently utilised, affected peak emission (see Figure 2.7 for Mark II luminaire basis functions), until the luminaires had 'burned-in', reaching thermal equilibrium at 75 minutes of operation.

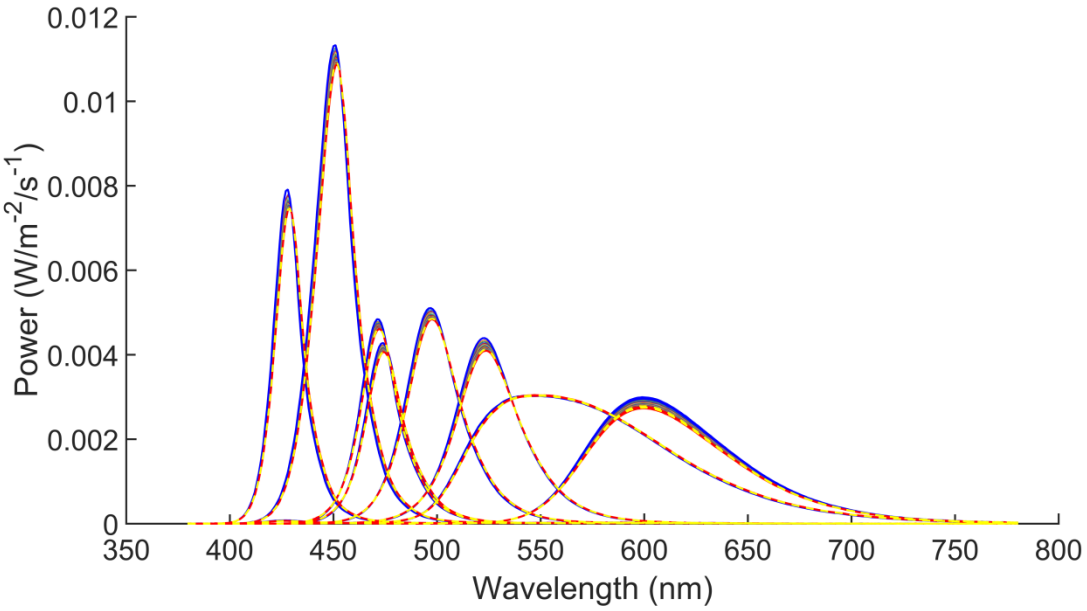

Figure 2.7. Basis functions (8 unique primaries) of the Mark II IREC luminaries, exhibiting linear peak shift. Blue lines indicate emission from cold start-up; red lines indicate basis functions after 3 hours of operation; dashed-yellow lines indicate emission after 75 minutes.

As each peak shift is linear for each function, knowledge of any one LED's peak emission used to derive the correct basis functions for colorimetric computations, as described below.

Spectral and Colorimetric Fitting

The known basis functions allow an accurate prediction of the colour and luminance of any produced spectrum. Given that a known spectrum's tristimulus value, T (XYZ), is the pointwise multiplication of that spectrum, S, and the colour matching functions, R, and that the spectrum produced from the LED lightsource is quantified by the linear combinations of its basis functions, A,

multiplied by a vector of weights,$\vec{w}$, it follows that the tristimulus value of any spectrum specified by $\vec{w}$ can be determined using Formula 2.7.

$$T = \vec{w} \cdot \boldsymbol{A}(\lambda) \cdot \boldsymbol{R}(\lambda) \cdot 683$$

Formula 2.7. Tristimulus value T is defined as the dot product of the spectrum (the multiplication of elements of scalar w, and the basis functions A), and the colour matching functions R, multiplied by the luminosity constant; which, is ~683 when radiance is measured in the SI unit $W \cdot m^{-2} \cdot s^{-1} \cdot nm^{-1}$ . .

Any spectrum that falls within the area under the curve of the basis functions can be matched colorimetrically by the LEDs. The illumination $A\vec{w}_i$, that would match or be the closest match spectrally to a target spectrum, B, can be solved for or at least optimised using linear least-squares fitting (see Formula 2.8), that is, by minimising the distance between the functions A and B with respect to wavelength.

$$\mathbf{A}\vec{w_i} = \boldsymbol{B}\ subject\ to: 0 \leq w_i \geq 1$$

$$\underset{w_i}{\text{Min}} \|\mathbf{B} - \mathbf{A}\vec{w}_i\|$$

Formula 2.8. Linear least-squares estimate, minimising the distance between the basis functions A and the target spectrum B, where the weights, w_i, are between 0 and 1.

As previously stated, there might be no perfect spectral match; however, a metamer of a desired spectrum can be created, where we do not specify the target spectrum's shape, but the tristimulus values only; $\vec{w}$ is solved by reducing the problem to three factors (see Formula 2.7). There is usually more than one metameric solution, a metamer set (Finlayson & Morovic, 2005), for each tristimulus value. One must now impose some arbitrary criteria for selecting a particular metamer from the set of possible solutions. In this work, we were predominantly interested in broadband illuminations, which are smooth and continuous, as defined by standard illuminations D65 and A (Schanda & International Commission on Illumination., 2007); therefore, the smoothest metamer was defined as a secondary characteristic of any optimisation as described by Finlayson, Mackiewicz, Hurlbert, Pearce and Crichton, (2014). The smoothness of a spectrum can be defined as the norm, the sums of the change in power with respect to wavelength, which is illustrated below and which is the same as taking the sum of the norm of the elements, squared, which can be applied using matrix D in Formula 2.9.

$$\int_l \left(\frac{dE}{d\lambda}\right)^2 d\lambda \approx \|\mathbf{DA}\overrightarrow{w_i}\|^2$$

$$\text{where } \mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix},$$

Formula 2.9. Formula for calculating the smoothness of a spectrum with respect to wavelength, and an equivalent expression using matrix multiplication from

Finlayson, Mackiewicz, Hurlbert, Pearce and Crichton, (2014); where $\frac{dE}{d\lambda}$ represents the gradient of the power of the spectrum with respect to wavelength, and D is the matrix which calculates the Euclidean norm of $\mathbf{A}\vec{w}$.

This can be posed as a quadratic programming problem, with the norm of the spectrum as the linear constraint, as formulated below (see Formula 2.10); examples of fits to standard illumination D65 using least-squares fitting and quadratic programming can be seen in Figure 2.11.

$$\min_{\vec{w}}(\vec{w_i}^{\boldsymbol{T}}\mathbf{A}^{\boldsymbol{T}}\mathbf{D}^{\boldsymbol{T}}\mathbf{D}\mathbf{A}\vec{w_i})$$

$$\text{subject to: } \mathbf{R}^{\boldsymbol{T}}\mathbf{A}\vec{w_i}\cdot 683 = \boldsymbol{T}$$

Formula 2.10. Expression as a quadratic equation to find the weights, $\vec{w}$, which maximises the smoothness of A constrained to the desired tristimulus value T. The variables are duplicated and transposed to make the computation one of matrix multiplication, such that it can be computed faster.

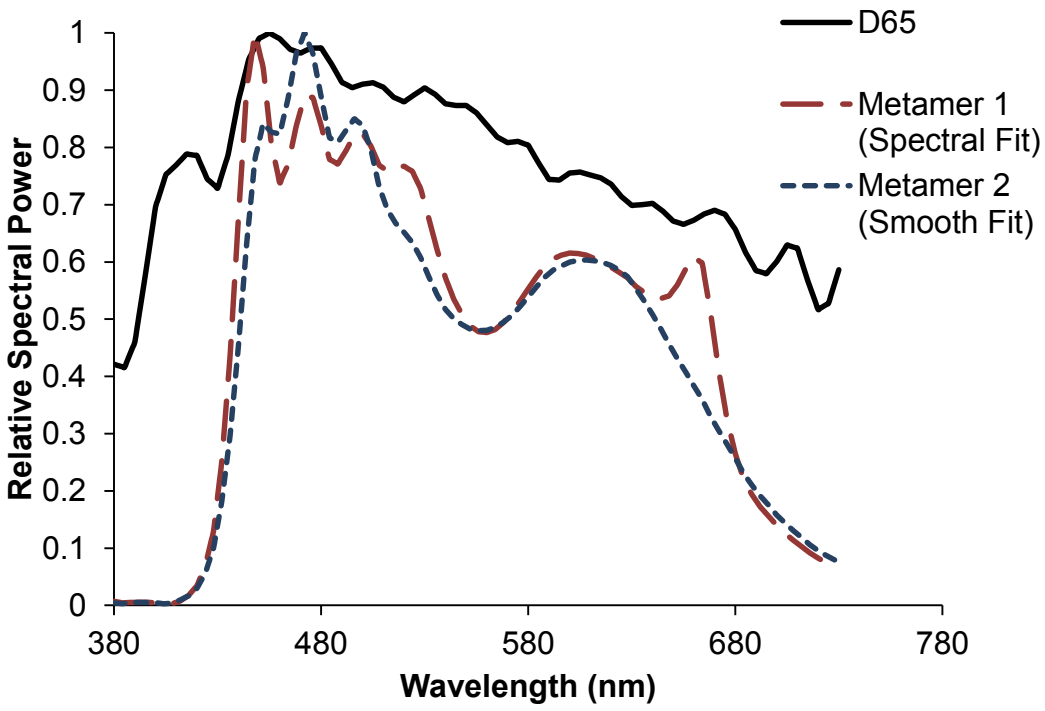

Figure 2.11. Standard illumination D65 and two matches: least-squares fitting and an additional constraint of smoothness using quadratic programming. Metamers produced for the IREC prototype luminaires, using basis functions gathered using the calibration methods detailed above.

The RS5B lightsources, which undergo considerable peak shift as the drive current increases, required the appropriate basis functions for the operating power to be selected for each metamer fit; otherwise the output spectra of the lamps did not match the theoretical spectra $\mathbf{A}\vec{w}$. This was performed by running the fitting code and then selecting the basis functions $\mathbf{A}_{n_i}$ where I was the band (10%-100%) that was closest to required $\vec{w}_n$, and then re-running the optimisation to get a closer, more accurate fit. The IREC prototype PWM luminaries did not suffer from this peak shift as described above and therefore

no further optimisation was required to find an initial weights vector. As operating temperature was a factor in spectral emission, fits were made using the 'burn-in' basis functions, and the lamps were allowed to reach thermal equilibrium before psychophysical experiments.

Creating real surfaces with specific chromaticities

As some of the experiments documented in this work used real scenes, there arose a need to produce surfaces (papers) with controlled, arbitrary chromaticities. This was achieved by characterising a desktop colour-printer, such that the output chromaticities on a printed page were predicted by modelling input RGB values (Bala, 2003; Ling, 2005).

The subtractive colour mixing process produces interactions between the conventional CMYK dyes used in desktop-printers which is much more non-linear than the additive colour mixing used in monitor display technology (Bala, 2003); as such, as the number of measured samples increases, so does the accuracy of any interpolation between those known chromaticity coordinates. However, typical desktop dye-cartridges are capable of only a few hundred printed patches before they require replacement or refilling, which in turn requires further calibration.

To overcome the limitations traditionally associated with printer calibration, the ink system was modified so that many samples could be printed (see Figure 2.12). As the number of measured patches was increased, measuring

one at a time would have been labour intensive, so a hyperspectral camera was used to measure many patches at once.

Figure 2.12. Photograph of HP Desktop printer with modified ink system.

Methods

Apparatus and Materials

A HP Photosmart desktop C310a, the printer to be calibrated, was operated by a Windows 7 computer, using MATLAB and custom driver software. Samples were printed on HP-branded thick, white, matte paper; 37 pages in total, 20 samples per page, 740 patches total, dimensions: ~3cm x ~2.5cm, and specified by random RGB values. A FotoRite Continuous-ink-system (CIS) was fitted inside the printer's 4 printing heads, with tubes that carried ink to the heads from 4 external reservoirs (see Figure 2.12); as described by (Smith, Söderbärg, & Björkengren, 1994). Each reservoir contained 500ml of dye (Cyan, Magenta,

Yellow or Black). A Gilden Photonics NVIR Spectral Camera was used to capture hyperspectral images, controlled by a Windows XP desktop computer and proprietary control software. A spectrally tuneable LED illuminator as described previously was used to light the samples, programmed by custom MATLAB control software to produce a smooth D65 metamer, again described earlier. A Formazin calibration tile produced by Gamma Scientific was placed within a Verivide viewing box and was used to measure the intensity of the D65 illumination across the entire captured scene.

Design and Procedure

The independent variables were the three 8-bit RGB values passed to the printer firmware, and the dependent variable was the spectra reflected from a printed patch of colour, under D65 illumination, and as measured by the hyperspectral camera at 4nm resolution between 400nm and 780nm.

The calibration tile was placed within the viewing box and the D65 illumination was set on the illuminator such that uniform illumination was produced within the viewing box. The camera was situated ~1m away from the scene, and focused such that only the calibration tile was in view, and a dark-image was taken whereby the camera-shutter was closed to measure background radiation. The shutter was then opened, and a 1920x800 HD image of the calibration tile was taken, where the spectra at each point of the image was ~D65 spectra at various intensities. Each of the 37 A4 pages of printed samples were placed against the calibration tile in turn and an image was taken of each

sheet (each referred to as a test image); a dark image was taken directly beforehand; see Figure 2.13 for scene example.

Figure 2.13. Rasterisation of a hyperspectral image, of a printed test sheet under uniform illumination; black squares mark the area over which spectra were averaged.

Analysis and Modelling

The hyperspectral images were imported into MATLAB. Firstly each element of the corresponding dark image was subtracted from each test image to adjust for background radiation. Secondly, the spectrum at each pixel in the test image was divided by the spectrum in the corresponding pixel of the calibration tile image, leaving the surface reflectance function of the imaged surface at that pixel. The mean of each patch was then taken on a wavelength-by-wavelength basis to produce a smooth reflectance function for each of the 740 patches.

As only RGB values can be specified to the printer, each patch's CIE XYZ tristimulus value was calculated from its SRF. A tetrahedral interpolation method was then used to create pyramids whose faces could be used to determine the amount of R, G and B to be used for any particular desired XYZ. The geometric path within that space required to achieve a desired chromaticity coordinate yields the contribution of each primary R G and B, as described in detail by Ling (2005); a full description of the mathematics are described by (Vrhel & Trussell, 2002).

These values can then be passed to the printer to create calibrated images. Half of the chromaticities measured were used in the model and half were used to test the model by comparing the actual chromaticities and the model's predicted values.

Chapter 3:

Colour constancy by illumination matching

Traditional colour constancy experiments, as reviewed in Chapter 1, have focussed on isolating the cues arising from raw cone-inputs, and the statistical properties of those inputs with respect to the scene, and how those cues can inform colour constancy mechanisms (Foster, 2011). They have used either highly controlled uniform backgrounds (Kraft & Brainard, 1999) or Mondrian scenes (Arend et al., 1991), usually on simulated displays, focussing little on the complex objects that exist in our natural environments. Little is known about the high-level cues relayed to the visual system by familiar objects, such a banana, or by the illuminations to which we are typical exposed, such as daylight.

None of the reviewed experiments (see Chapter 1), have focused on systematically testing the contribution of real, familiar (e.g. a banana) and chromatically-matched, novel objects (e.g. a yellow cube) on colour constancy. Moreover, most of the reviewed experiments use a small number of illuminations with a high degree of perceptual distance between them (Kanematsu & Brainard, 2013); none have focused on the colour constancy for both the small and large illumination changes that are experienced ecologically. Moreover, nearly all colour constancy experiments have changed the contrast of surfaces within a scene by either adjustment of test patches or manipulation of surface statistics (Foster, 2011; Smithson, 2005). This means that the upper limits of colour constancy have not been measured; that is, when the scene is held unchanged and the illumination changes, what are the conditions under which perfect constancy is obtained, and what overall is the level of observers' colour constancy?

In the experiments described here, the assumption is that if the illumination on a scene changes, and the observer cannot detect the change, then the observer is colour constant. That is, colour constancy mechanisms have compensated for the change in the illumination, and the observer cannot consciously discern a change. If an observer notices the illumination has changed, colour constancy can be said to be incomplete. Note that in this case, it would be the case that operationally, the observer is colour-constant, in being able to attribute the change to a change in the illumination and not the surface colours (in the definition of operational colour constancy as proposed by Craven and Foster (1992)) but here the definition is of appearance constancy; there is a change in appearance, and therefore constancy is not perfect. The point at which an observer discerns an illumination change, just above chance, is their threshold for the physical change in the illumination at which colour constancy mechanisms fail.

Here, a new method of colour constancy is used to quantify colour constancy – illumination matching. Participants' discrimination thresholds for systematically controlled illumination changes are established. Participants viewed a scene, in a viewing box, illuminated by a 'target' illumination in each trial, then two illuminations were presented sequentially, each separated by a dark period; one of the comparisons was always the target, another was some perceptual distance ($\Delta E_{u^*v^*}$) away from the target. Observers were asked to match which of the two alternatives matched the target light. Based on the central assumption discussed above, the principle of this task is simple – if the

observer is colour constant, then they will not be able to select reliably above chance which of the alternatives is the target, as both appear the same.

This allows for very small illumination changes to be compared with large illumination changes, and for natural changes, such as those typical to daylight, to be compared with unnatural illumination changes of the same perceptual distance, such as, a green illumination change that differs by the same perceptual amount as a blue daylight change. The hypothesis that colour constancy mechanisms perform better for daylight illuminations was tested by generating two sets of metamers: one set of metamers of daylight illuminations that varied along the "blue-yellow" axis near the Planckian locus, and a distinct set of metamers of the same size and varying in equal steps along a "green-red" axis, which crossed the centre of the daylight locus at D67 (6700 K). The prediction was that if colour constancy was better for daylight illumination changes then matching accuracy would be poorer for these bluer and yellower illuminations, at equal perceptual distances, than their green-red counterparts. It is worth noting that $\Delta E_{u^*v^*}$ change does not guarantee equal changes in scene statistics, which are dependent on scene contents and the consistency of the spectral shape between illumination changes.

The hypothesis that colour constancy is better for familiar objects was tested by having observers complete the illumination matching experiment when the viewing box contained either: natural objects (an apple, banana and realistic fake pear); or paper primitives (a paper cube, cylinder and pyramid), matched in colour and luminance to the surface reflectance average of the fruits. The

prediction was that if colour constancy was better for familiar objects, then the illumination matching accuracy would be poorer for all illumination changes when the scenes contained those familiar objects.

Finally, the hypothesis that a grey surface could cue colour constancy mechanisms (Foster, 2011) was tested using two scene backgrounds. The viewing box was lined with either grey card, or a variegated (Mondrian) background (containing no achromatic patches). If colour constancy was indeed better for the grey background, then matching accuracy would be poorer for all light changes in conditions using the grey background.

Experiment 1.1

Methods

Ethics and declaration

This work has been presented in Pearce, Crichton, Mackiewicz, Finlayson and Hurlbert (2014), and is presented here in accordance with the CC BY licence. The experiment was conducted in accordance with both the Faculty of Medical Sciences (FMS) and APA Ethics guidelines. The experiment was granted ethics approval (reference number 00312) by Newcastle University's FMS Ethics Committee. Written consent from each participant was obtained after a complete disclosure of the procedure was outlined.

Participants

Eight naïve observers (6 female; aged 20-28yrs mean 26yrs) participated in the study. All participants were recruited using opportunity sampling on a first-

come, first-serve basis through the Institute of Neuroscience Research Volunteer Program. All participants were healthy with normal visual acuity, or wore corrective eyewear. Participants were screened for colour deficiency using the Ishihara Plates and the Farnsworth-Munsell 100-Hue Test (Mean TES: 24, Kinnear & Sahraie, (2002); all participants had normal colour vision).

Design

A two-alternative forced-choice 2x2x3 (illumination change axis, by box lining, by box contents) repeated-measures design was utilised, using a method of constant stimuli. The independent variables were the illumination set used (either the blue-yellow daylight locus or red-green orthogonal locus illuminations, see Figure 3.1) and the contents of the viewing box which was lined with either grey or Mondrian card (see Figure 3.2), consisting of either no objects, fruits or novel objects. The dependent variable was the measured accuracy for illumination matching for each element of each illumination set. This resulted in 6 conditions in total, two background conditions by 3 box-contents conditions.

Apparatus

A spectrally tuneable illuminator was used, as described in Chapter 2 and pictured in Figure 3.2, containing 6 RS5B lightsources, each with a bank of 10, independently-addressable LED channels, which projected into an integrating sphere that reflected the light almost perfectly-uniformly into the viewing box also described earlier (see Chapter 2). An XBOX-360 gaming pad was used to interface with the control software, written in MATLAB running on a Windows 7

desktop computer, which also controlled the illuminator. Headphones were used to produce low-latency audio using an ASIO enabled sound card.

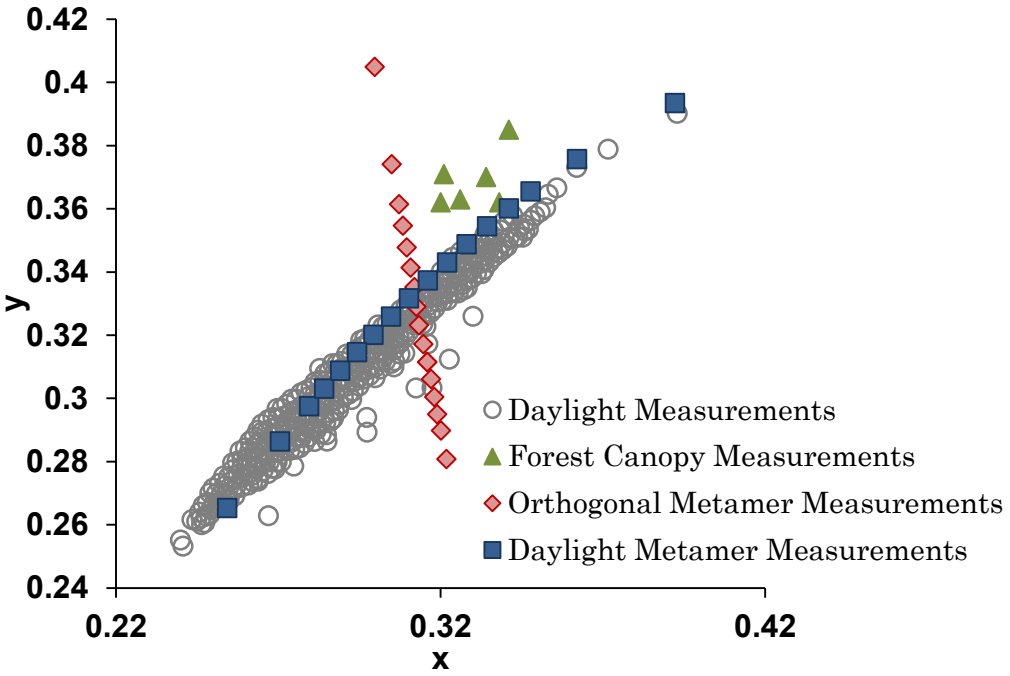

Figure 3.1. Generated metamers of daylight illuminations and novel, orthogonal illuminations that cross at D67, atop daylight measurements made by Hernandez-Andres et al., (2001). Green markers show the most extreme greenish illuminations measured by Sumner and Mollon (2000).

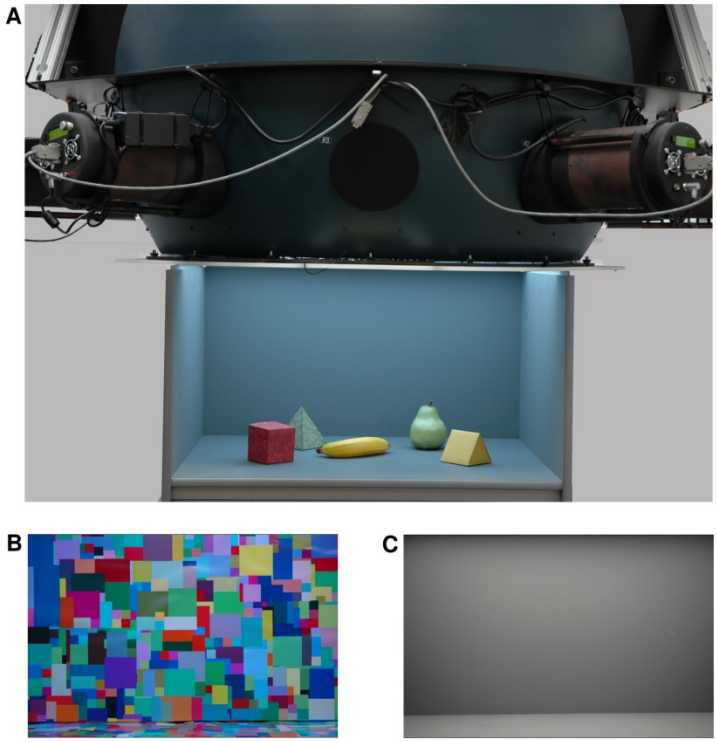

Figure 3.2. A: Photograph of the spectrally tuneable illuminator, illuminating the viewing box (front removed) and containing fruits and chromatically matched novel objects. *B:* Mondrian box lining used inside the viewing box for Mondrian background conditions. *C:* Grey card lining used for grey background conditions. From Pearce, Crichton, Mackiewicz, Finlayson and Hurlbert (2014).

Stimuli

The viewing box walls and floor were lined with either grey card (mean CIE 1931 coordinates x = 0.299, y = 0.324, under the D67 illumination), or Mondrian card (x = 0.321,y = 0.359, under D67; see Figure 3.2, C & B respectively). The Mondrian card contained patches varying in size (~0.2cm to ~ 12.0cm, ~7.6 deg. of vis. ang. at a viewing distance of 90cm). The viewing box also containing either (1) no objects, (2) 3D fruit objects (an apple, banana or fake pear), or (3) novel 3D

primitives made from paper printed with colours matching the fruits in mean and range of chromaticities; a cube which matched the apple, a triangular prism with the banana and a pyramid which matched the pear (see Appendix 2 for tabulated chromaticities). Printing was carried out with the calibrated printer described in Chapter 2; with the exception of the Mondrian card, which was printed using a wide-format, inkjet, sRGB calibrated printer (through a Newcastle University Printing Service), using random non-achromatic sRGB values for each patch.

Two sets of illuminations, each consisting of 17 distinct spectra chosen to be metameric to particular chromaticities sampled along either (1) the daylight locus, from blue to yellow, and (2) an orthogonal locus, red to green, were generated for presentation with the illuminator (using the techniques described in Chapter 2). (The distinct spectra are termed "metamers" for short.) Two target illuminations on each locus, ± 10 perceptual steps $\Delta E_{u^*v^*}$ from D67, and 11 comparison illuminations: ± 0, 6, 12, 18, 24 and 58 $\Delta E_{u^*v^*}$, from each target illumination were generated (see Figure 3.1). Measurements of extreme blue and yellow comparisons (58 $\Delta E_{u^*v^*}$ away from target) can be seen in Appendix 3. All generated metamers had a luminance of between 22.49 and 23.85 cd/m^2 as measured from the white calibration tile in the viewing box; moreover, the full metamer set showed mean change from the generated chromaticities of only 1.19 $\Delta E_{u^*v^*}$ over the 6 week testing period.

Procedure

Participants were seated before the viewing box and asked to look inside, through the viewing aperture. Their heads were not fixed, but the minimum distance of ~90cm from the visible scene was constrained by the box front. The scene was not visible before the experiment began as no illumination was produced by the illuminator. Participants were given standardised instructions which directed them to two buttons on the Xbox gamepad, 1 and 2. The instructions read: "*You will be shown a light that illuminates the viewing box; this is the target light. Then there will be two subsequent lights, you are asked to signal which is most like the target light, using either of the buttons, [1] denoting the first light is most similar, or [2] for the second light*" (Pearce et al, 2014). Participants then completed a 2-minute dark adaptation period before the start of the experiment.

Each trial began with 3 audible tones, generated by the computer and delivered by headphones; this signalled a new trial to the observer. The illuminator, synced with the auditory cue, illuminated the viewing box with one of the target illuminations for 2000ms, selected at random. The illuminator was then switched off for 400ms before a further auditory tone was presented, signalling the first comparison illumination. The first comparison illumination was then presented, and illuminated the viewing box for 1000ms. A further 400ms dark-period followed, before a tone signalling the second and final comparison. The second comparison was then presented for 1000ms. One of the comparison illuminations was always the target illumination, while the other differed by some perceptual distance (± 0, 6, 12, 18, 24 and 58 $\Delta E_{u^*v^*}$). After the

two comparisons were shown, a dark-period followed a tone that cued the participant to respond, answering which comparison illumination matched the target. Trials were self-paced with the next trial following the observer's response, with a minimum of 1000ms dark period between each trial. Each participant completed 480 trials per condition (2880 in total over the 6 conditions; with 240 trials per locus, two targets per locus, and 10 per comparison). Participants completed a mandatory break after every 120 trials, and were informed that they could break or withdraw at any time.

Each experimental condition, 6 in total, was conducted in a separate session; each session was ~60 minutes in length. All sessions were conducted over a 6 week period, at the participant's convenience.

Control Experiment

One of the comparison illuminations, the most extreme 'red' comparison (+58 $\Delta E_{u^*v^*}$ on the orthogonal locus; 10 trials per participant), was not presented correctly as there was a miscommunication between the control software and the illuminator; this resulted in the +0 $\Delta E_{u^*v^*}$ comparison (identical to the target) being shown 10 more times. Participants therefore performed not significantly different from chance on these trials. To confirm the level of matching accuracy for this comparison, a control experiment with 4 participants was conducted, using the grey viewing box with no objects. Accuracy for the extreme red comparison illumination in this control experiment was not significantly different from the other ±58 ΔE comparisons, and not significantly different from 100%.

Results

Participants could complete the task with mean matching accuracy of μ = 72.47% across all comparison illuminations and conditions. A repeated measured ANOVA with three independent variables revealed that there was a significant difference in accuracy between the daylight and the orthogonal loci ($F(1,7) = 17.404$ $p < .01$); with mean matching accuracy (percent correct) lower for the daylight illumination changes (70.20% vs 74.74%). Mean accuracy for the grey background condition (μ = 76.37%) was significantly higher than for the Mondrian background (μ = 68.57%; $F(1,7) = 11.385$, $p = .012$) (Figure 3.3). No significant difference was observed for the conditions containing different objects: fruit, novel objects or no objects ($F(2,6) = 1.776$, $p = .248$).

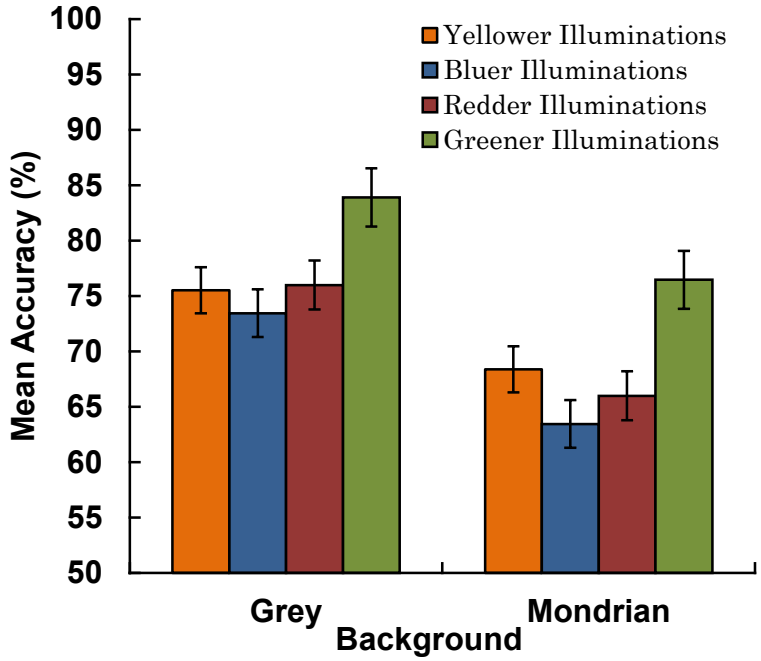

Figure 3.3. Mean matching accuracy for all observers across all conditions for the four illumination directions in scene containing either a Mondrian or grey background. Error bars indicate ±1 SEM.

For finer analysis of the illumination matching patterns, each locus was divided into two parts by splitting each locus at the central point (D67), thereby creating four loci of chromatic change directions: bluer, redder, greener, and yellower illuminations. A subsequent repeated-measured ANOVA with Greenhouse-Geisser correction shows that over all conditions, mean accuracy was significantly different for the four colour directions ($F(2.12,14.85) = 15.031$, $p < .01$; Figure 3.4). Illumination matching was poorest for bluer illuminations and best for greener illuminations; post-hoc tests using Tukey's HSD revealed all chromatic change directions were significantly different from one other ($p < .05$) for the grey background and Mondrian conditions. Each illumination direction was also significantly different from each other, except for red and yellow illuminations, which did not differ significantly ($p < .05$).

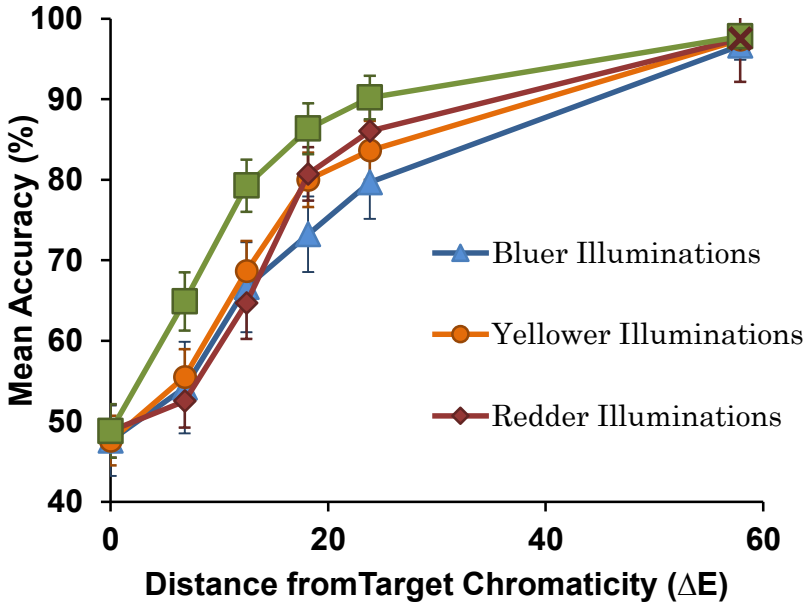

Figure 3.4. Mean accuracy for all observers in all conditions for each chromatic direction as a function of ΔE from the target illumination. Error bars indicate ±1 SEM. Red-cross indicated control experiment data (n= 4).

Discussion

In the surface matching experiment paradigm, colour constancy requires the illumination to be discarded by the visual system, so that matches to a target surface are close to the surface reflectance function of that surface; as if it were illuminated by a white illumination. If the illumination changes, and the visual system is perfectly colour constant, surfaces will appear the same, yielding the same matches as produced under a neutral illumination. Therefore, colour constancy can be measured by illumination matching, with poor discrimination of an illumination change (an inability to match the correct illumination) signaling high colour constancy; that is, surfaces signaling the illumination appear the same under all comparisons.

Illumination discrimination was poorest for blue illumination changes, signaling the best colour constancy for these light changes. Colour constancy was poorest for greener illumination changes, as demonstrated by the highest illumination discrimination accuracy at each ΔE step.

The results demonstrate a clear difference between chromatic change directions. Bluish vs. yellowish changes produced different discrimination accuracies, as did reddish vs. greenish changes. These asymmetries were not expected on the basis of the "blue-yellow" and "red-green" colour-opponent channels that are initiated by cone-contrast processing in the retina. Bluish daylight illumination changes, which appear to be experienced most frequently (based on inspection of measurements reported by Hernandez-Andres et al., 2001) were discriminated least reliably in this experiment, supporting the

hypothesis that colour constancy should be better for daylight illuminations. Other studies have shown better colour constancy for blue-yellow than red-green illumination changes (Worthey, 1985); however this is the first demonstration that discrimination is specifically enhanced in the greenish illumination change direction.

Contrary to the hypothesis that familiar objects might cue mechanisms of colour constancy, performance was no higher for familiar vs. novel objects, with matched surface chromaticities. This result supports the evidence presented by Kanematsu & Brainard (2013). This might be due to the cues present in the background overweighing the information presented by the objects.

Colour constancy was, though, improved for the Mondrian scene over the grey scene for all illumination changes. The difference in performance between the Mondrian and grey scenes suggests that the presence of a larger number of distinct surfaces aids colour constancy mechanisms. This might be due to scene articulation (Linnell & Foster, 2002). The reasoning is that illumination changes become less discriminable when scene articulation increases, not that more information about the illumination *per se* becomes available; this reasoning is also supported by the performance for the grey scene being higher, as changes may be more obvious when a single uniform surface changes. Nevertheless, the bias for particular illumination directions is preserved in both scenes suggesting a universal bias for bluer illumination directions that spans scene contents.

As classical theories of colour constancy (grey-world and max-flux) do not allow for illumination bias (Finlayson, Hordley, & Hubel, 2001), they do not

predict the observed bias for illumination direction, and may be considered incomplete in this regard. This said, it is crucial to know what is the signal change from each surface in each scene for each illumination change, to predict whether: a) the scene average in both scenes is biased as to predict the observed performance; b) the brightest point in each scene is the same, providing a consistent cue to the illumination change; and c) the cone-contrast coordinates of the change for each scene are symmetrical, to determine whether the bias arises after retinal computation.

To test these three hypotheses, hyperspectral images were taken of each viewing box background, Mondrian and grey, under each of the comparison illuminations and one of the target illuminations. The cone-contrast signal change was computed at each point in the hyperspectral image between the target and each comparison to precisely quantify the signal available to the observer to perform the task. Thus it was expected that if the bias for particular illuminations could be explained by scene information, this would be available in the images. Moreover, if the bias occurred via processing at higher stages in the visual pathway, symmetrical signal change would be observed for the blue-yellow and red-green cone-contrast channels. Finally, the difference in performance between the grey and Mondrian scenes with respect to the grey-world and max-flux hypothesis was tested to determine whether the mean or maximum change (at any single location) might cue the visual system to the illumination change.

Experiment 1.2

Methods

Apparatus and Materials

The hyperspectral camera, as described in Chapter 2, was used to capture an image of the viewing box, described earlier, containing either the Mondrian or grey card lining. The illuminator, programmed with the target and comparison illuminations, described earlier, was used to illuminate the viewing box for each image. All apparatus was controlled using custom software written in C++ and Matlab, and proprietary software, on Windows 7, and Windows XP workstations.

Procedure

The camera was positioned at ~90cm (±1cm) away from the viewing box, and focussed such that the scene was clearly in view. The illuminator was programmed with one of the illuminations. The hyperspectral camera lens was closed and a dark-image was taken to allow for background noise. Then the lens was opened and an image of the viewing box was taken. This process was replicated for each of the comparison illuminations. The images were each stored as a 1900x800x96 element array, measuring the spectra from 400nm to 780nm at 4nm resolution for each of the 1,520,000 pixels in the image. These images were loaded into MATLAB for subsequent analysis.

Analysis

Each image had its corresponding dark-image subtracted to reduce the contribution of noise to the image. Then 95 patches on the Mondrian scene were chosen, and the same corresponding coordinates were taken for the grey scene.

The spectra from those points were averaged to obtain the mean reflectance from each surface and the chromaticity coordinates in the CIE 1931 colour space were calculated for the spectra. The perceptual distance, $\Delta E_{u^*v^*}$, and cone-contrast coordinates were then calculated between the target (reference) scene and the comparison scene using the MacLeod-Boynton cone-contrast space (MCB; see Equation 3.5); this obtained the signal change at each point. The mean and maximum changes for each cone-channel (Luminance, BY and RG) were then calculated, along with the distribution statistics of these changes, such as the skewness and kurtosis of those distributions.

Red-green contrast: $L\text{-}M = 1953 * (r^{patch}_{McB} - r^{white}_{McB})$

Blue-yellow contrast: $S\text{-}(L+M) = 5533 * (b^{patch}_{McB} - b^{white}_{McB})$

Luminance contrast: $L+M = (lum^{patch}_{McB} - lum^{white}_{McB}) / lum^{white}_{McB}$

Equation 3.5. Modified Macleod-Boynton transformation as proposed in McDermott and Webster (2012); each MCB was computed as the scaled McB coordinates of the stimulus relative (in contrast) to the McB coordinates of the target illumination whitepoint. The McB coordinates r_{McB} , b_{McB}, and lum_{McB} are defined as $l/(l+m)$, $s/(l+m)$ and $(l+m)$ respectively, where l, m and s are the long-, middle- and short-wavelength cone excitations of the stimulus.

Results

The scene average change for the Mondrian and grey scenes was approximately equivalent to the illumination change in all cases, and an independent samples t-test revealed that the distributions of changes were not

significantly different from each other $t(42) = .083$, $p = .775$. However, the whitepoint of each box was not centered on D67, as can be seen in Figure 3.6. Moreover, the distribution of cone-contrast changes for the Mondrian box were symmetrical; an example of the MCB change for the BY channel can be seen in Figure 3.7. Maximum luminance change was < 0.5 cd/m^2 between the illuminations, with the skewness and kurtosis of the distributions for blue and yellow illumination changes and red and green illumination changes being symmetrical in all cases: MCB Lum 0.06 and -0.06 for blue and yellow illuminations, with kurtosis of 6 and a skewness of -2 and +2 respectively. Each distribution of MCB changes for yellower and bluer, redder and greener illumination changes cause a magnitude of change in one direction of each chromatic axis (redder-greener, bluer-yellower) relative to the target illumination. The distributions for blue and yellow, and red and green were collapsed together to create 2 distributions, to test if they were indeed both tails of each distribution were normal and indeed both sides of single normal distributions. The Shapiro-Wilk test of normality indicated that both Blue-Yellow distributions (S-W = .179, df = 9, p = .20) and Red-Green distributions (S-W 1.99, df = 9, p = .20) were normal.

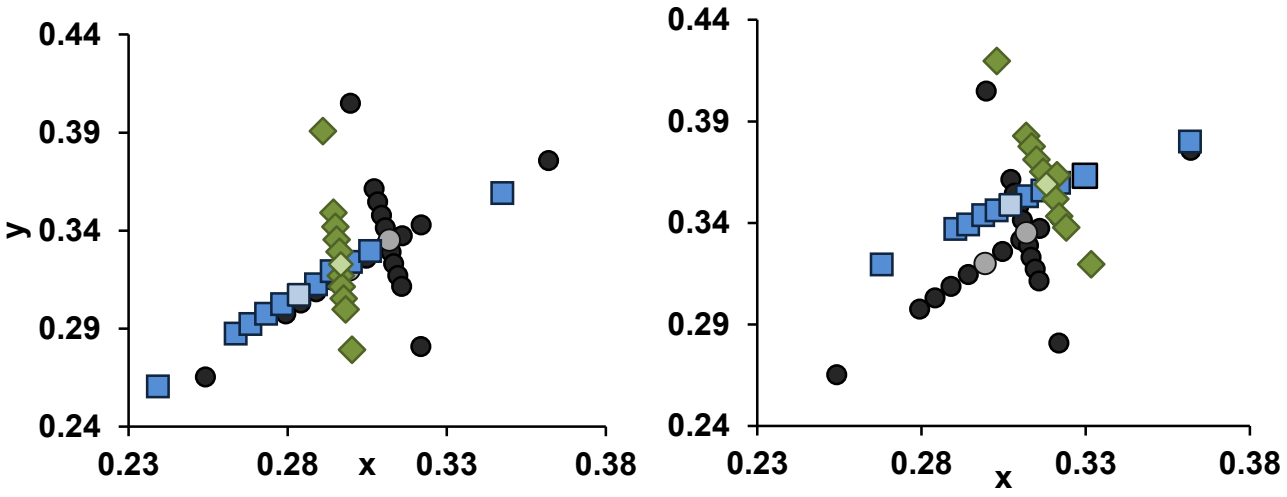

Figure 3.6. Box averages of Mondrain (right) and Grey (left) lined viewing box; from the hyperspectral images, in CIE 1931 colour space. Light markers indicate target illuminations, grey markers are the illumination chromaticities and blue and green markers are daylight and orthogonal chromaticities, respectively. Black circles mark the illumination chromaticities of exit spectra from the integrating sphere.

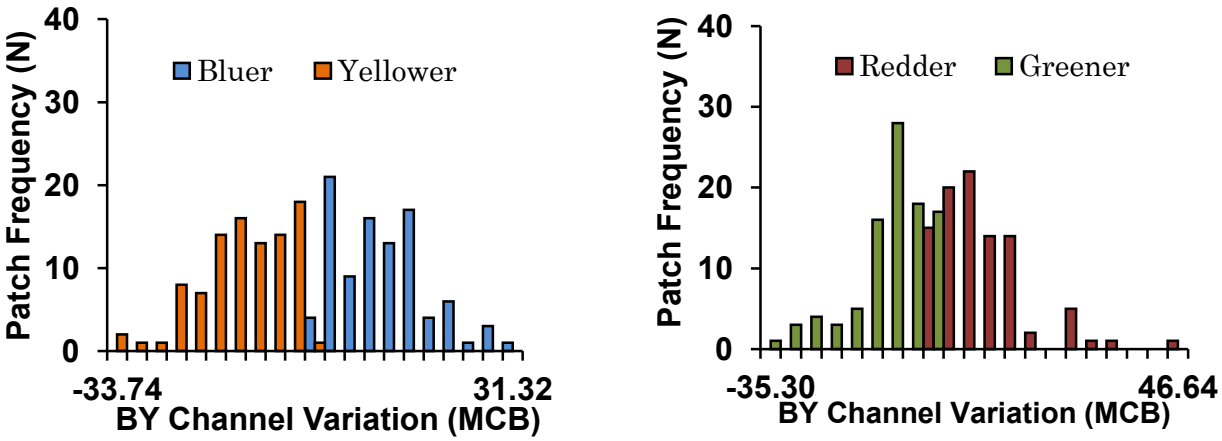

Figure 3.7. Frequency of each of the 95 patches corresponding to levels of BY cone contrast change for each of the bluer, yellower, redder and greener comparison illumination changes at 18 $\Delta E_{u^*v^*}$ away from the target illumination.

Discussion

The shape and magnitude of signal change for each scene does not explain the differences in illumination discrimination performance between the chromatic directions. In each instance, these changes were symmetrical between the cone-contrast axes (MCB: Lum/BY/RG).

The hyperspectral images revealed that neither scene was completely neutral on average, with the Mondrian box being slightly greenish and the grey box being slightly bluish on average. However, the bias in illumination discrimination is preserved despite the changes in scene average. This suggests that the changes in the scene average might cue colour constancy mechanisms and explain absolute performance levels; nonetheless, these differences in scene average cannot explain the bias between illumination directions. To confirm that the changes in box average were not skewing the performance statistic the scene average ΔE change was used instead of the illumination change ΔE in a further ANOVA, and performance is still significantly greater for the grey background ($F(1,29) = 51.692$, $p < .001$).

The brightest patch was always the same in the grey scene, as the whole surface changed by the same amount. However, in the Mondrian scene, the brightest patch differed between the illumination directions. Maximum luminance change in each scene did not predict performance; for example, maximum change for yellower illuminations does not correlate with mean performance for those illumination changes, but does correlate highly with performance for greener illuminations ($r = .884$, $p < .05$); moreover, the maximum luminance change for bluer illuminations correlates with performance

on all but redder illuminations ($r = .979$, $.960$, $.977$; $p < .05$, for yellow, blue and green illumination changes respectively). Therefore, max-flux cannot account for either overall or direction specific illumination discrimination performance.

As can be seen in Figure 3.7, the blue-yellow and red-green channel changes are equally distributed, as is the overall scene change. These magnitudes should represent the change in the physical signal; however, the fact that the observer's response is asymmetric to equivalent changes suggests that the retinal signal may not fully account for the observer's responses. The observer is perceptually blind to the illumination change (i.e. colour constant), but to differing degrees depending on the colour of the illumination change. The results lend weight to the notion that higher-level cortical processes are responsible for the observed bias and that the visual system is optimized for the natural environment. For example, dichoptic presentation of scenes under different illuminants results in a mixed state of adaptation, showing that chromatic adaptation is at some level completed cortically (Werner, Sharpe, & Zrenner, 2000).

The bias towards blue illuminations, without a bias in the physical scene, could be shaped by an expectation of the visual system that a blue change in a scene is due to the illumination, as is evident in natural images (Webster et al., 2007) . A visual system that optimizes colour constancy to typical changes (blue daylight) and not when an atypical change has occurred (green illumination change) gains an advantage by directing attention only when meaningful variations in the environment are present.

Chapter 4:

Illumination and Chromatic Colour Discrimination.

In the previous experiment, there was a clear asymmetry in performance accuracy of the participants' illumination discrimination for each chromatic direction, with performance for blue illuminations being systematically lower. Asymmetries in chromatic discrimination for colour patch matching tasks are well described by discrimination ellipses (Krauskopf & Gegenfurtner, 1992; MacAdam, 1942; Wyszecki & Stiles, 1982). These ellipses illustrate the just-noticeable difference from a central colour at each point along the ellipse contour, with all chromaticities falling inside the ellipse being indistinguishable to the observer. Figure 4.1 shows the threshold for each chromatic direction, which was linearly interpolated from the 75% point, from mean participant data from Experiment 1.1; it is plotted atop the MacAdam ellipse for D65, centred on D67 (there is no ellipse for D67, however D67 sits on the D65 boundary).

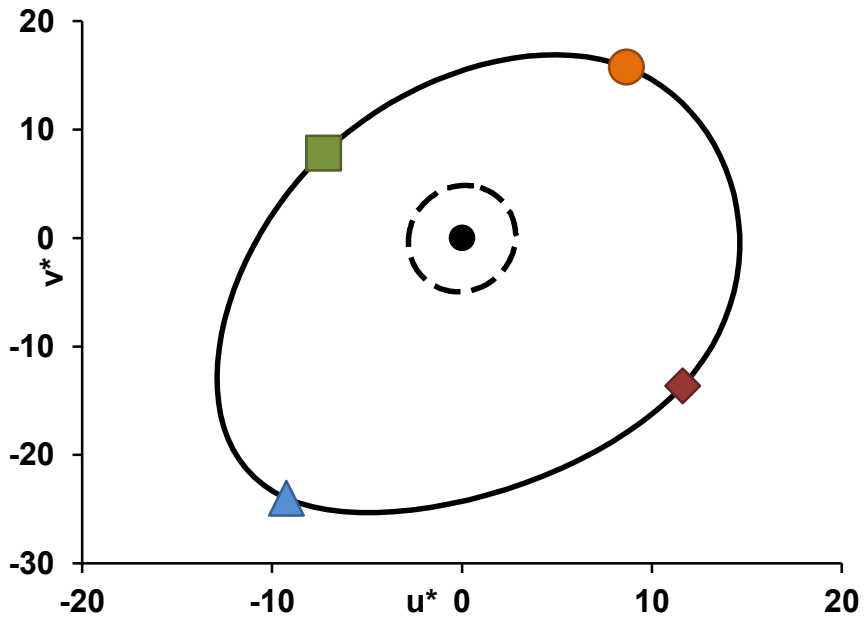

Figure 4.1. Spline of illumination discrimination thresholds plotted in CIE Lu*v* colour space; atop MacAdam ellipse for D65 (dotted line), cantered on D67 (spot).

The MacAdam ellipse, plotted in Lu*v* colour space is roughly circular, as the space is primarily constructed to be perceptually uniform; based on MacAdam ellipse judgements (Schanda & International Commission on Illumination., 2007). As can be seen, the thresholds for blue-yellow and red-green illumination discrimination are much larger and asymmetric than the corresponding values represented by the MacAdam ellipse. The general magnitude of the difference between the MacAdam ellipse and the illumination discrimination ellipse is partly due to the difference in task; the illuminations are presented successively rather than simultaneously, and the discrimination is global rather than local. Thresholds for surface matches have been shown to change in magnitude depending on the adaptation point but generally are assumed to be symmetrical between opponent directions (Krauskopf & Gegenfurtner, 1992). This suggests that illumination discrimination may be different to surface colour discrimination, and in turn suggests that the $\Delta E_{u^*v^*}$ model does not hold for the illumination discrimination task.

In the previous experiment the 'method of constant stimuli' design was employed. Each participant was exposed to each illumination contrast the same number of times, maximising the likelihood of finding a bias between the chromatic directions as the *n* for each contrast is high; however, using this method, each observer's absolute threshold for illumination change is not directly established.

The aim in the following experiment was to determine whether there is any evidence that an asymmetry exists between the two chromatic directions

generally, that is, between the red-green and blue-yellow cone-opponent axis; therefore, thresholds were established for observer's illumination and surface discrimination ability, using comparable chromaticities. To do so, a different method of establishing thresholds for both illumination and surface discrimination was utilised so that discrimination thresholds could be established directly.

To establish the observer's threshold of just-noticeable difference for illumination discrimination, an adaptive staircase method was employed. Instead of presenting a set of fixed contrasts, illuminations between 0 $\Delta E_{u^*v^*}$ and 50 $\Delta E_{u^*v^*}$ from D67 in each chromatic colour direction were generated. Observers completed the same task as before, and were asked to pick which of two illuminations matched the target, D67. The one differing comparison was chosen by an interleaved adaptive staircase which stepped up and down until the observer was guessing 1 or 2, at which point the observer's threshold was established. It was expected that the threshold illumination difference for changes in the blue direction (called "blue thresholds" for short) would be the highest as colour constancy mechanisms would inhibit change detection for these illuminations; conversely, it was expected that green thresholds would be the lowest as colour constancy mechanisms work poorest for these illumination changes.

The second task was completed on a computer monitor. In this task, observers viewed a grey field (D67), and were asked to indicate in each trial if an arrow, presented briefly for 500ms, pointed left or right. The arrow was of

uniform chromaticity and appeared ~2° horizontal and ~1.5° vertical in visual angle. The arrow's chromaticity differed from the background in cone-contrast units of either ±BY or ±RG; the arrow was either: redder, greener, yellower or bluer than the background but remained the same luminance. The level of cone-contrast was also controlled by an interleaved, adaptive staircase procedure. It was predicted that the task would establish the absolute threshold for cone-contrast discrimination for each observer. Therefore, it was predicted that if illumination discrimination thresholds were indeed different from surface discrimination $\Delta E_{u^*v^*}$ thresholds, for each opponent axis, then the asymmetries in the illumination discrimination task arise from a bias rather than a lack of sensitivity.

Experiment 2.1 –Cone-Contrast Discrimination Test (CCDT)

Methods

Ethics and declaration

The Cone-Contrast Discrimination Test (CCDT) is described in further detail elsewhere (Cranwell, Pearce, Loveridge, & Hurlbert, 2014; submitted). The experiment was conducted in accordance with guidelines and with approval of the Committee of the Faculty of Medical Sciences at Newcastle University (reference numbers: 00312, 00612). All participants gave informed, written consent before the start of each of the two detailed experiments, and were

debriefed fully afterwards. All participants completed both experiment 2.1 and 2.2.

Participants

Six observers (4 male, mean age 22 y; range 21-23) participated in the study. All participants were student researchers at the Institute of Neuroscience, Newcastle University. All participants had normal visual acuity and did not have anomalies of colour vision. Participants were recruited by personal request of the experimenter and were all experts in colour vision research. .

Design

A two-alternative, forced-choice procedure utilising two, interleaved, adaptive staircases was used for each of two conditions. The two independent variables were: the pointing direction of a constant-sized arrow that was presented in the centre of a computer monitor, pointing either left or right, and the cone-contrast coordinates of that arrow with respect to a D67 grey background; the arrow chromaticity was controlled by the interleaved staircases, which modulated a different chromatic axis in each condition. The staircase was a 1 up and 2 down staircase, with 6 different step sizes (25, 20, 15, 10, 3, 1 steps). The dependent variable was the measured cone-contrast coordinates needed between the arrow and the background for the observer to detect which direction the arrow pointed.

Apparatus

A black viewing tunnel fitted with a chin rest was used to position each participant's head ~77cm from an ASUS 10-bit addressable LCD (PA238Q) computer monitor. The monitor was driven by a Windows 7 computer equipped with a PNY 600 10-bit graphics card. Custom computer software written using the Psychtoolbox library for MATLAB (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) controlled the monitor and experimental procedure. Participants gave feedback using a computer mouse.

A PR650 spectroradiometer was used to calibrate the computer monitor, as described by Brainard (1995). The exit spectra were collected over 1° of visual angel from ~1m away, for 1024 steps (between fully open and fully closed), for each thin-film transistor channel in the display matrix. This produced a 10-bit look-up table for three independent channels; linear combinations were used to produce each desired chromaticity with fixed luminance. The mean between the measured and modelled chromaticities was $\mu\Delta E_{u^*v^*} = 0.22$.

Stimuli

A D67 grey field (Yxy = 64, .3145, .3388) which filled the monitor screen, ~36˚ of horizontal and ~21˚ vertical visual angle, and was present throughout the experiment. An arrow of uniform chromaticity was presented in each trial, which pointed either left or right, ~2˚ of horizontal, and ~1.5˚ of vertical visual angle, and whose chromaticity was modulated in four colour directions: red, green, blue and yellow on the two cone axes S-(L+M), in condition 1, and (L-M) in condition 2; the maximum contrast was ~25 $\Delta E_{u^*v^*}$ from the background, which corresponded to cone-contrast units in each direction indexed in nominal units 1

through 75; these are not documented here, as they change per monitor, per calibration; however, each step corresponded to ~0.314 $\Delta E_{u^*v^*}$ contrast. A white fixation square, approximately 0.5 ° in visual angle was presented in each trial (Yxy = 250, .3181, .3393).

Procedure

Participants were given standard instructions, which can be found in Appendix 4. Participants were asked to place their heads on the chin rest and told to fixate on the white fixation square at the beginning of each trial. There was a maximum of 50 trials for each of the two interleaved staircases, for each of the two conditions: chromatic axis S-(L+M) and (L-M); participants completed a maximum of 200 trials or until each of the staircase procedures had completed, having reversed more than 6 times on a single chromaticity or 30 times total.

Trials were self-paced; a fixation square was presented for 1000ms before a 500ms blank period, with just the D67 background visible; the arrow was then presented, pointing left or right determined at random on a trial by trial basis; the arrow was positioned randomly ±5.5° of visual angle both horizontally and vertically from the central fixation square. The arrow was presented for 150ms, after which a further blank, where only the D67 background was present, until the participant responded. Participants responded using the computer mouse, pressing either the left mouse button for a left-pointing arrow, or the right button for the right-pointing arrow. A correct response was relayed to the staircase procedure which then reduced the contrast level between the arrow and the background, by the current step size, for the next trial in that staircase; an

incorrect response would trigger a reversal where the contrast would be increased. All staircases started at maximum contrast (~25 $\Delta E_{u^*v^*}$) and stepped towards minimum contrast (0 $\Delta E_{u^*v^*}$). The staircase used for the trial was determined at random at the start of each trial. The entire procedure, for both axes, took ~15 minutes.

Thresholds

To determine each observer's threshold for each axis of change, each nominal reversal of the staircase was converted to the corresponding cone-contrast value via a look-up table. The last 4 reversals were then averaged to obtain a cone-contrast threshold, which was then converted to CIE $\Delta E_{u^*v^*}$ with respect to the background for comparisons between colour axis.

Experiment 2.2 – Illumination Discrimination

Methods

Design

A two-alternative forced choice adaptive staircase procedure was used in a 1x4 repeated-measures design. The independent variable was the distance of a comparison illumination in $\Delta E_{u^*v^*}$ colour space away from a central illumination (D67 CIE Yxy = 200 .31 .32). The dependent variable was the participant's $\Delta E_{u^*v^*}$ threshold for illumination change computed by the average of reversals on each of three adaptive staircases for the four levels of the IV; the illumination directions: bluer, redder, yellower and greener. Staircases were 2 up, and 1

down, with 5 step levels (15, 10, 5, 3 and 1), targeting observer's 70.71 correct point (García-Pérez, 1998)

Apparatus

A spectrally tuneable luminaire system was used for the main experiment, with 12 LED primaries (Ledmotive Prototype light sources). The viewing box was also as used previously, with dimensions 71cm (width) x 77cm (depth) x 47cm (height) and a viewing aperture of 7.5cm height and 14.5cm width; lined with grey card, as used previously. An XBOX 360 game controller was used for participants to give feedback, using the Windows driver software and custom software written in MATLAB. Headphones were used to give audio cues to the state of the trials and were controlled by a low-latency ASIO sound card. All hardware was controlled using custom software written in MATLAB.

The illuminator was calibrated using methods described in Chapter 2.

Stimuli

The viewing box was lined with Grey Card (mean x = 0.299, y = 0.324, under D67), which was identical to that used in previous experiments. Four sets of illuminations were used. Each set had 51 elements, with illuminations ranging from 0 to 50 $\Delta E_{u^*v^*}$ away from a central point D67. The 4 illumination sets were bluer and yellower illuminations whose chromaticity was determined by their distance from D67 in $\Delta E_{u^*v^*}$ along the Planckian Locus, transformed from CIE Yxy 1931 colour space, and redder and greener illuminations which were defined by their distance from D67 in $\Delta E_{u^*v^*}$ along the isotemperature-line 6700 CCT.

Illumination Generation

Illuminations were generated using the quadratic programming method of fitting described by Mackiewicz, Pearce, Crichton, Finlayson and Hurlbert (2014), and detailed in the metamer generation section of Chapter 2. Metamers were generated with the constraint that the sum of the linear combinations of the basis functions multiplied by the weights of each functions produces the smoothest continuous function, that is, the flattest spectrum for that chromaticity. All illuminations generated were isoluminant with a luminance of 250 cd/m^2. All theoretically generated illuminations were verified using the PR650 spectroradiometer; these spectra are plotted in Figure 4.2.

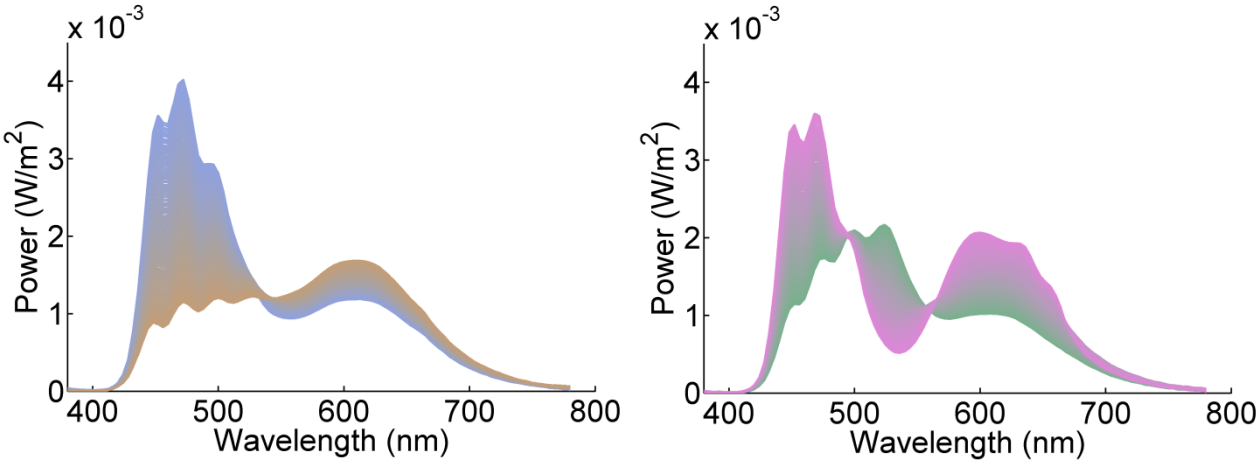

Figure 4.2. Left: spectra of the blue and yellow metamers, ranging from extreme blue to extreme yellow. Right: spectra of red and green metamers, ranging from both extremes. Both sets share a common central point, D67.

Procedure

The general procedure described here is the same as Experiment 1.1, and varies only in the way comparisons were selected, described here for completeness.

Consenting participants, after reading standardised instructions, were asked to sit in front of the viewing box aperture as close as possible so that they could see into the viewing box; the head was not fixed. Once participants were situated comfortably, participants were given a 2 minute dark adaptation period before the main procedure.

Trials were self-paced, with the experimental procedure ending when either the limit of 300 trials for all staircases or each having reached the limit of 6 reversals. Participants were asked to match which of two comparison illuminations matched a target illumination shown in each trial, by pressing either of 2 marked buttons of the game controller pad; 1 for the first comparison illumination presented, and 2 for the second. Trials started with 2 audible tones synchronised to the presentation of the target illumination that lit the viewing box for 2000ms. A dark period of 400ms separated the two comparison illuminations that were then shown sequentially, each for 1000ms, synchronised to audible tones. A tone was sounded 400ms after the second comparison illumination to cue participant response. The target illumination, and one of the comparison illuminations (position assigned randomly) was D67; the other comparison illumination was determined by the adaptive staircase, and perceptually differed between 0 and 50 $\Delta E_{u^*v^*}$ away from D67; starting at a

random position between 20 and 40. If the participant correctly matched the D67 comparison to the target, a correct response was sent to the staircase procedure which stepped down closer to target illumination in perceptual distance. If an incorrect response was sent, the staircase would reverse and step up in the opposite direction. The staircases within and between illumination colour directions were interleaved, and were selected randomly until all had completed. Each participant completed 12 staircases, 3 for each of the 4 illumination colour directions. A break was given every 100 trials, followed by a further dark adaptation period.

Obtaining Thresholds

The nominal values of each step in the staircase were converted to their actual $\Delta E_{u^*v^*}$ values using a look-up table, and the last 4 reversals of each staircase were averaged to obtain a just-discriminable threshold determined by that illumination staircase, from the D67 target illumination. The mean of three thresholds per staircases, per illumination direction were taken to get the observer's just-noticeable difference for that direction in CIE $\Delta E_{u^*v^*}$ units.

Results

Mean $\Delta E_{u^*v^*}$ thresholds, for the CCDT, in the four colour directions, (green: 3.11, red: 2.53, blue: 3.77 and yellow: 3.26) were not significantly different from each other $F(3,15) = 1.842$, $p = .183$; with a mean contrast threshold of 3.17 $\Delta E_{u^*v^*}$ required for chromatic target detection. Mean thresholds for the illumination discrimination task were significantly different from each other $F(3,15) = 7.672$, $p = .002$; means can be seen in Figure 4.3, which shows the four

chromatic illumination directions: redder, greener, bluer and yellower – against the four colour directions for surface matching: bluer (+BY), yellower (-BY), greener (-RG) and redder (+RG). Post-hoc pairwise comparisons revealed a significant difference between green illumination thresholds and each other colour direction (red: p = .005, blue: p = .017 and yellow: p = .042). There was a significant difference between thresholds between the two tasks $F(1,5) = 17.852$, p = .008.

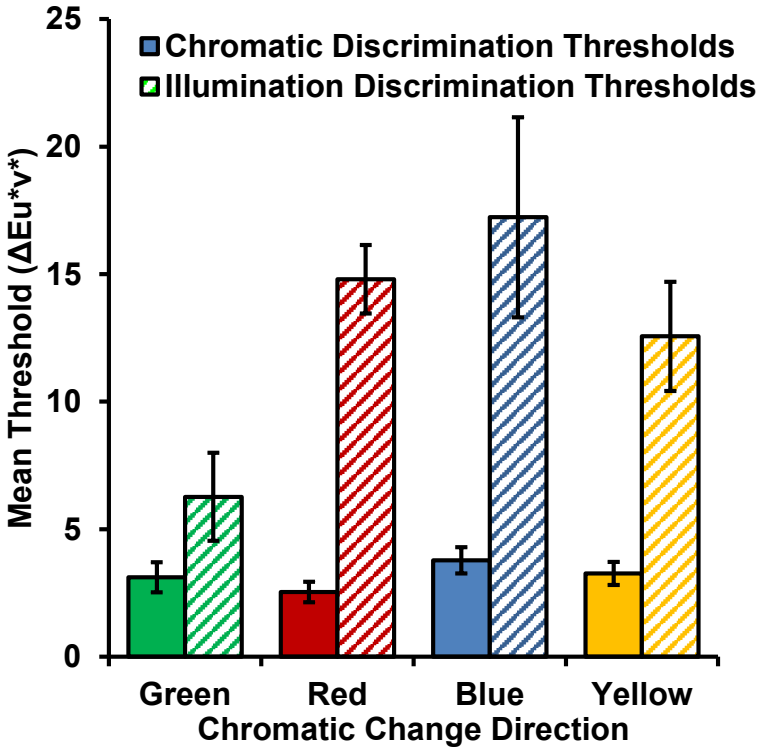

Figure 4.3. Illumination discrimination thresholds (Line Fill) against chromatic discrimination thresholds (solid fill) for each chromatic direction: blue, red, green, yellow. Error bars show ±1 SEM (n = 6).

Discussion

As predicted, illumination discrimination thresholds produce the same asymmetry for each chromatic direction as in Experiment 1.1. This asymmetry is not present for chromatic discrimination as seen in the CCDT task; which suggests that the observer's sensitivity for illumination changes of equal magnitude (which displays asymmetries) does not reflect solely the symmetry of the low-level processes. The general magnitude of the chromatic discrimination thresholds ($\Delta E_{u^*v^*}$ ~3 rather than 1, implied by the units) is most likely due to the task, as the arrow is flickered only briefly; this is consistent with thresholds being higher and symmetric across each chromatic direction, with task difficulty affecting each chromatic direction equally. It may also be true that the measured illumination discrimination thresholds are higher than the observer's absolute threshold due to the task, however, no other experiments test systematic illumination discrimination; therefore, a benchmark (such as 1 perceptual step in $\Delta E_{u^*v^*}$) does not exist to make testable predictions. However, the difference between green thresholds in both tasks suggests that inflation in thresholds, due to task, may be roughly equal in both instances, because green illumination discrimination is very similar to observer's measured green chromatic discrimination.

In terms of signal, a larger field of change is present in the illumination discrimination experiment; however, the entire scene undergoes the same transformation; whereas, in the chromatic discrimination task the target changes locally with respect to the background. However, the similarity between the green illumination and chromatic discrimination thresholds suggests that

any spatial component does not affect each chromatic channel equally. Indeed, contrast threshold differences have been shown for S-cone changes and L/M-cone changes as a function of spatial frequency (Cao, Zele, Smith, & Pokorny, 2008), however, better chromatic discrimination for blue changes around an adaptation point is expected (Krauskopf & Gegenfurtner, 1992); suggesting that illumination discrimination should be best for the blue-yellow contrast axis around D67, which is not observed.

Performance differences on the two tasks could be explained by the violation of the relationship between surface chromaticities, as proposed by relational colour constancy (Craven & Foster, 1992), cuing the visual system to a surface colour change; however, when all surfaces change similarly, as in the illumination discrimination task, colour constancy mechanisms inhibit the change, interpreting it as one due to the illumination.

Chapter 5:

The effects of Achromatic Surfaces on Colour Constancy

In Experiment 1.1, there was a systematic difference between illumination discrimination thresholds for the grey and Mondrian lined box; with accuracy significantly higher (colour constancy significantly poorer) for the grey box lining conditions. It is not clear whether colour constancy was poorer for this condition because: a) there were not adequate comparison surfaces present in the scene to signal the likely surface reflectance function of the box lining's and thus, the upper limit of achievable colour constancy was lower; or, b) that the presence of a large uniform surface, a reference surface, allowed the illumination change to be identified easier than in the Mondrian scene, where the luminance and contrast of some patches would change less than others during each illumination change. As illumination discrimination does not examine the subjective experience of the colours within a scene, it cannot determine how the sensation of the grey box lining was under the illumination changes; however, it can be determined whether levels of colour constancy are modulated when a reference surface is added to the scene.

The concept of reference surface is common among computational algorithms, which usually attempt to use such surfaces to determine the illumination spectrum or chromaticity; Retinex for example, assumes that a surface reflecting the illumination most accurately within a scene will be a surface of maximal reflectance under that given illumination (Hurlbert, 1989; Land, 1977), and normalises the image to this 'white'. Such an approach requires only identifying pixels with maximal reflectance, while other approaches try to identify surfaces of known reflectance to determine the illumination (Gijsenij, Gevers, & van de Weijer, 2011; Novak, 1991). If the visual

system has a representation of a reference surface within a scene, with an assumed reflectance, it could use such a surface to determine the direction and magnitude of an illumination change (Brainard & Maloney, 2011). Such reference surfaces must first be identified, perhaps from other properties of the object (Vurro, Ling, & Hurlbert, 2013), or due to that surface's distribution of surface chromaticities under illumination changes (Ling & Hurlbert, 2008a). Ling and Hurlbert (2008a) demonstrated that the chromaticity distributions of natural objects were predictable under changing illuminations and may be used by the visual system to increase colour constancy. Vurro, Ling and Hurlbert (2013) demonstrated how surface properties which are invariant with chromatic change in the illumination, shape and texture for example, increase memory colour for natural objects. Furthermore, evidence from Crichton, Pichat, Mackiewicz, Tian and Hurlbert (2012) demonstrates that skin chromaticity gamuts transform predictably under natural illuminations, and that these transformations can be used with gamut-mapping algorithms to retrieve the illumination incident upon the scene. The aforementioned authors were able to adequately predict the cone-contrast change due to an illumination change, by determining a set of diagonal matrices that translates the gamut shift of human skin under those illumination changes. This suggests that perhaps natural objects like skin could be used by the visual system to determine the colour of the illumination. Of course, the source of the distribution of chromaticities is not important to the gamut-mapping algorithm (Crichton, Pichat, et al., 2012); any object transforming in a similar way under illumination changes could be used to

retrieve the illumination; for example, a grey uniform surface as suggested by Foster (2011).

It may be the case therefore, that colour of the natural objects used in experiment 1.1 remained largely colour constant, and that the cues used for the illumination discrimination task were solely derived from the surrounding scene; however, specific surfaces such as skin, surfaces bearing skin chromaticities, or the addition of large uniform, achromatic surfaces to a scene could cue the visual system to the illumination. Participants completed the illumination discrimination task, as described previously, for two types of illumination changes, in 4 scenes. The illumination changes were either blue daylight changes, or green novel changes, as previously described. The scenes each contained the variegated Mondrian lining and contained either: a prosthetic hand, whose average chromaticity matched that of the Caucasian skin average chromaticity described by (Crichton et al., 2012); a wood block, whose chromaticity average was the same as the hand; or a grey block whose chromaticity matched that of the grey box lining, described previously.

It was predicted that if colour constancy mechanisms perform better for scenes containing surfaces with a broadband reflectance function (surfaces that predictably transform with the illumination chromaticity), then discrimination thresholds would be higher for scenes containing these objects. Instead, if these objects make the illumination change more obvious, then thresholds for illumination discrimination would be lower for these scenes. Furthermore, if colour constancy mechanisms are additionally informed by stored knowledge

about the transformation of surface chromaticities of familiar objects, then it is predicted that illumination discrimination thresholds would be higher for the scene containing the prosthetic hand than for the chromaticity-matched block. Finally, it was predicted that the asymmetry between green and blue discrimination thresholds would be preserved with better colour constancy observed for blue illumination changes.

Experiment 3.1

Methods

Participants

Ten naïve observers (5 female; mean age 20 y; range 19-24) participated in the study. All participants were recruited using the Institute of Neuroscience Research Volunteer Program. All participants had normal or corrected to normal visual acuity and were screened for colour deficiency using the Ishihara Colour Plates, and the CCDT (described in Experiment 2.1); 1 participant, not documented here, was excluded as they were identified as having a colour vision deficiency. Participants were paid £7 per session and each attended 2 sessions. The experiment was reviewed and conducted in accordance with guidelines set by the Ethics Committee of the Faculty of Medical Sciences at Newcastle University (reference number 00312); all participants gave informed, written consent before the start of the experiment.

Design

A two-alternative forced choice adaptive staircase procedure was used in a 2x4 repeated-measures design. The independent variables were the distance of comparison illuminations in $\Delta E_{u^*v^*}$ colour space away from a central illumination (D67 CIE Yxy = 200 .31 .32), and the contents of a viewing box lined with Mondrian paper: either empty, with a prosthetic hand, flesh-coloured block or a grey block. The dependent variable was the participant's $\Delta E_{u^*v^*}$ threshold for illumination change computed by the average of reversals on each of three adaptive staircases; for two illumination directions, blue and green. Staircases were 2 up, and 1 down, with 5 step levels (15, 10, 5, 3 and 1); with average $\Delta E_{u^*v^*}$ of 1.88 per step away from the target illumination.

Stimuli

The viewing box was lined with Mondrian Card (mean x = 0.321, y = 0.359, under D67), which was identical to that used in previous experiments. In each of four conditions the box contained either: no objects, a prosthetic hand or a grey or flesh painted block of wood; see Figure 5.1. The prosthetic hand was from a Simulaids' manikin, which was painted with acrylic paints to match chromaticities from hyperspectral-images of skin under D67 illumination (mean x =.410, y=.390), taken by Crichton, Pichat, Mackiewicz, Tian and Hurlbert (2012). The flesh coloured block was painted with the same paint as the hand, to have the same mean chromaticity under D67. The grey block was painted with grey paint to match the grey viewing box in the previous experiment, with mean chromaticity (x = 0.299, y = 0.324) under D67 illumination. Two sets of illuminations were used. Each set had 51 elements, with illuminations ranging

from 0 to 50 $\Delta E_{u^*v^*}$ away from a central point D67. The 2 illumination sets were bluer illuminations whose chromaticity were determined by their distance from D67 in $\Delta E_{u^*v^*}$ along the Planckian Locus, transformed from CIE Yxy 1931 colour space, and greener illuminations which were defined by their distance from D67 in $\Delta E_{u^*v^*}$ along the isotemperature-line 6700 CCT; both as used in Experiment 1.1 & 2.2.

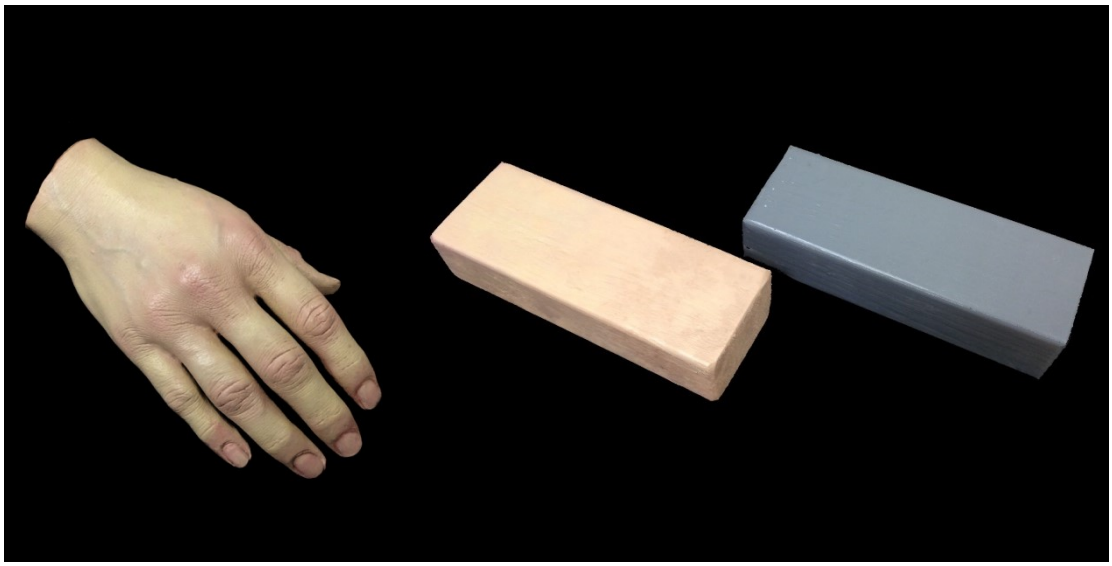

Figure 5.1. Photographs of stimuli used in scenes: Left: a prosthetic hand painted to have the same mean chromaticity of an average Caucasian hand described by Crichton et al. (2012), with tones matches from hyperspectral images. Right: flesh coloured block with same mean chromaticity and tones as hand, and grey block with same dimension, painted to match chromaticity of grey box in Experiment 1.1; colour differences are visible due to uncalibrated imaging and printing process.

Procedure

Participants attended two sessions. In each session the participants completed two conditions, which were decided pseudo-randomly. In either the first session or the second session participants completed the CCDT test (as described in experiment 2.1), decided by coin toss. All participants completed the Ishihara plates and the Farnsworth-Munsell test in the first session.

Consenting participants, after reading standardised instructions, were asked to sit in front of the viewing box aperture as close as possible so that they could see into the viewing box; the head was not fixed. In each experimental condition the viewing box contained different or no objects. Once participants were situated comfortably, participants were given a 2 minute dark adaptation period before the main procedure.

The main procedure was identical to experiment 1 and 2, with observers being asked to match which of two comparison lights matched a target (D67). The main difference was that in this experiment only blue and green comparison illuminations were used. Each participant completed 6 staircases in each of the four conditions, 3 for each of the 2 illumination colour directions. Staircases were interleaved within and between chromatic directions. A break was given every 100 trials, followed by a further dark adaptation period.

Obtaining Thresholds

Thresholds were obtained using the same method as in experiment 2.2, with nominal staircase values converted distance from D67 in $\Delta E_{u^*v^*}$ units, before averaging the last 5 reversal values to obtain a just-discriminable difference from the D67 target illumination.

Results

In each condition, thresholds for blue and green illumination changes were not normally distributed as indicated by the Kolmogrov-Smirnov test, as such parametric statistics were inappropriate for hypothesis testing; participants' mean blue and green thresholds over all conditions can be seen in Figure 5.2; the Wilcoxon signed ranks non-parametric test was used to test differences between illumination change direction and conditions, which use median scores, as can be seen in the box plots shown in Figure 5.3. The tests indicated that blue and green thresholds for illumination changes were significantly different ($z = -2.19$, $p = .168$; $z = -2.805$, $p = .03$; $z = -2.803$, $p = .03$; $z = -2.805$, $p = .03$) for each conditions baring the condition with the hand present (p value corrected for 6 comparisons). The Wilcoxon test indicated a significant difference between the no object and grey block conditions for blue illumination thresholds only ($z = -2.310$, $p = .021$), and a difference between the no-object condition and the flesh block was trending in significance ($z = -1.836$, $p = .066$).

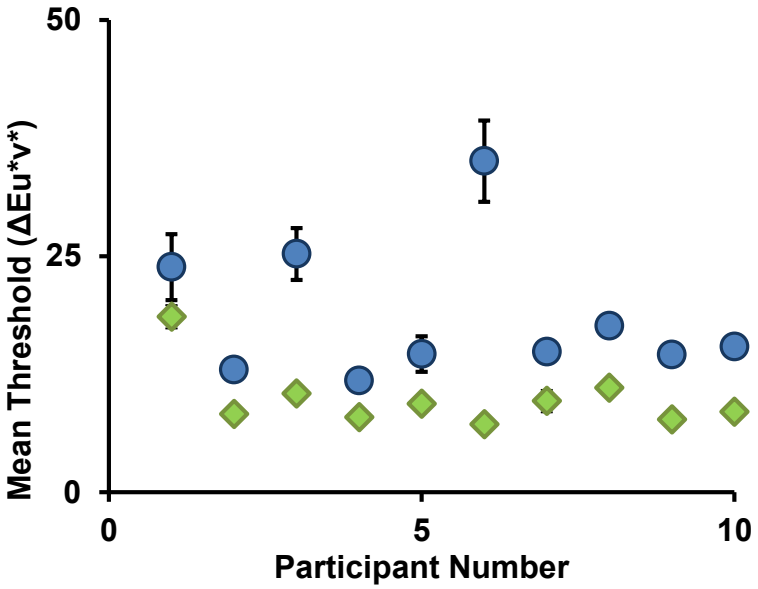

Figure 5.2. Participant's mean thresholds for blue (circles) and green (diamonds) illumination changes, across all conditions; error bars show ±1 SEM (n=10).

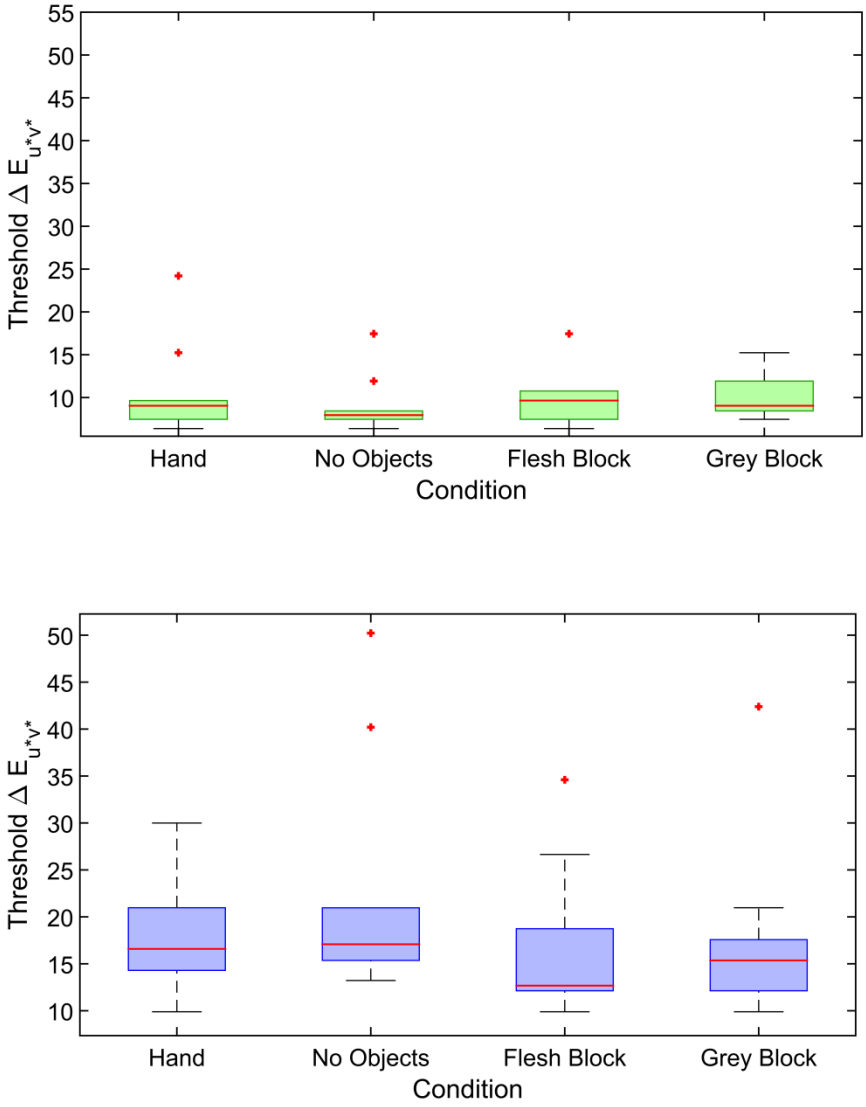

Figure 5.3. The median illumination discrimination thresholds for each of the four conditions, shown with a red line for each condition; green thresholds above and blue below. Red marks indicate outliers

Discussion

The results indicate that the presence of a hand or chromatically matched flesh coloured object neither facilitates colour constancy nor makes the illumination change more discriminable, contrary to the predictions from the computational literature which suggest that additional information about the

illumination chromaticity should be available from these cues (Crichton et al, 2012). Indeed, although there is a visible difference in the mean thresholds between these conditions, and it is in the expected direction for making an illumination change more discriminable (poorer constancy), it is not statistically significant. Thresholds for green illumination appear unaffected by the change in box contents, inclusive of the grey block condition; conversely, the presence of a grey object appears to make bluer illumination changes more discriminable, replicating the reduction in thresholds observed in previous experiments, for the grey box conditions.

A perfectly colour constant system should see a grey surface as the same under changing illuminations (Hurlbert, 1999; Rutherford & Brainard, 2000). Defining colour constancy as purely the stability of surface colours across changing illuminations leads to the conclusion that these data show colour constancy is strictly worse for conditions containing grey objects, immediately following an illumination change. On the other hand, defining colour constancy as the ability to attribute a change in the scene correctly to either an illumination or surface material change – operational colour constancy (Craven & Foster, 1992) -- leads to the opposite conclusion: operational constancy is facilitated by the presence of an achromatic surface, as the surface provides a reliable cue to the illumination change, meaning that this change in relation to the rest of the scene can cue the visual system to a change in illumination.

If a grey object is a reliable cue to the illumination, computational models of colour constancy would predict that chromatic adaptation would be better for

these scenes, as a grey surface would both a) move the scene average chromaticity closer to that of illumination and b) provide a reliable contrast change that is due to the illumination. These computational models are not temporally sensitive, and as demonstrated by Linnell and Foster (1996), observers are better at discriminating between illumination changes and surface reflectance change at lower latencies. These data are therefore consistent with the hypothesis that a neutral surface will cue the visual system to an illumination change, and over long periods will provide a reliable cue to the illumination (Foster, 2011; MacDonald & Roque, 2013). Stated differently, each cone observing a grey surface will adapt to a closer approximation of the illumination over time, and will exhibit a larger change in potential, on average, when the illumination changes.

While these data also demonstrate a bias for the blue illumination change direction, the task does not distinguish between an explicit illumination discrimination task, and a surface discrimination task; that is, participant's thresholds colour might be a measure of sensitivity to the colour change in the scene average, rather than an explicit illumination change threshold. To systematically measure the use of cues in the scene, and further elucidate the biases observed in illumination discrimination tasks, tighter control needs to be established over the observed scene and surface materials; such as using a computer simulation of these scenes, and systematically varying a single cue to the illumination.

Most importantly these data show that, when comparing colour constancy under different illumination changes, an achromatic surface should not be used within a scene, as it may artificially skew constancy indices for those illuminations. In fact, many constancy experiments have included grey surfaces while testing, e.g. the effects of surround contrast, global average and max-flux (Kraft & Brainard, 1999); the role of familiar objects in illumination discrimination (Granzier & Gegenfurtner, 2013; Pearce, Crichton, Mackiewicz, Finlayson & Hurlbert, 2013) and the role of familiar illuminations on colour constancy (Pearce, Crichton, et al., 2014; Xiao et al., 2012). The conclusions concerning the influence of these factors may be in need of revision when derived from experiments without a grey surface present.

Chapter 6:

Colour constancy through illumination Discrimination in real and simulated scenes.

The observed bias in illumination discrimination thresholds has thus far only been observed for real scenes, where there is a physical distinction between the illumination and the surface. In other colour constancy experiments, such as achromatic adjustments, levels of performance are established using a computer monitor (Smithson, 2005), where the distinction between illuminations and surfaces is purely simulated.

Various cues to the illumination are present in real scenes which cannot be controlled precisely or represented accurately in computer displays; such as, the information within the dynamic range in natural scenes, which may be many thousands times greater than on desktop displays (Xiao, DiCarlo, Catrysse & Wandell, 2002); specular highlights from glossy objects, which appear to affect constancy indices (Xiao et al., 2012; Yang & Maloney, 2001) or mutual reflections (Bloj, Kersten, & Hurlbert, 1999).

To control these possible extraneous variables, and determine if the biases for illumination discrimination thresholds also arise in simulated scenes, we generated a simulated Mondrian scene under simulated illuminations. The surface geometry of a Mondrian scene was modelled in Blender, and then passed to a physically based renderer, Mitsuba using Rendertoolbox 3 (Heasly, Cottaris, Lichtman, Xiao, & Brainard, 2014; Jakob, Arbree, Moon, Bala, & Marschner, 2010). The Mondrian scene was rendered under each of 50 illuminations for each comparison colour direction; using spectra measured from the illuminator used in experiment 2 onwards (see Figure 4.2).

Observers viewed the Mondrian scene under simulated D67 illumination, through a stereoscopic setup documented by Xiao, Hurst, MacIntyre and Brainard (2012), and performed the illumination matching task described earlier. It was expected that if an extraneous variable in the real scene was driving the observed bias in illumination discrimination thresholds, then participant thresholds would be flat, as in the chromatic discrimination task (CCDT); moreover, it was expected that if the distinction between surfaces and illuminations was not necessary for the task, as is predicted by other colour constancy experiments using a monitor, then the bias would be observed as in the real scene.

Experiment 4.1

Methods

Ethics

This work has been previously presented in Pearce et al. (2014); presented here under open licence. The experiment was conducted in accordance with APA ethical principles. Ethical approval was granted by ethics committees at both the University of Pennsylvania and Newcastle University (reference number 00312). Written consent was given by all participants before experimental sessions and after full disclosure of the procedure.

Participants

Twenty-one, healthy participants (aged between 19-26 yrs) participated in two experiments. Nine participants took part in the simulated condition at the

University of Pennsylvania (mean age 19.7yrs; 6 male), and 12 participants took part in the real-scene condition at Newcastle University (mean age 21.5, 8 male). Participants were recruited by each university's participant recruitment program, for undergraduate course credit. All participants had normal or corrected to normal visual acuity and normal colour vision as determined by the Ishihara Plates test.

Design

A two-alternative forced-choice, 4x2 repeated-measures, between-subjects design was utilised, using a 1-up, 2-down, adaptable, interleaved staircase procedure. The first independent variable was the illumination change direction, either: bluer, redder, greener or yellower from 0 to $50\Delta E_{u^*v^*}$ from a D67 standard illumination; the second independent variable was the scene condition, either: a real Mondrian scene, under real illumination changes; or a simulated Mondrian scene on a stereoscopic display, under simulated illuminations. The dependent variable was the established illumination discrimination thresholds of the observers to illumination changes, determined by the averaging of the last 5 reversals of the staircase procedure.

Apparatus

Three spectrally-tuneable luminaires, with 12 LED primaries (Ledmotive Mark I Prototype light sources) were used to illuminate the Mondrian viewing box – discussed previously. These luminaries were housed inside an illuminator room ($2m^3$ volume), painted with highly reflective white paint, into which observers viewed the Mondrian box, from outside, through a viewing aperture

(with dimensions as in Experiment 2), see Figure 6.1. A stereoscopic viewing apparatus was used, consisting of 2 LCD monitors which projected towards two mirrors, which in turn reflected light through viewing holes in a black wall, to observers – this setup is documented in detail by Xiao et al., (2012). A gamepad was used in both setups for observers to give feedback. The stereoscopic display, luminaires and staircase procedure were controlled using custom software and the MATLAB package. The luminaires were controlled by a Windows 7 computer, via Bluetooth, and the stereoscope was controlled by an Apple iMac computer, running OS X Leopard.

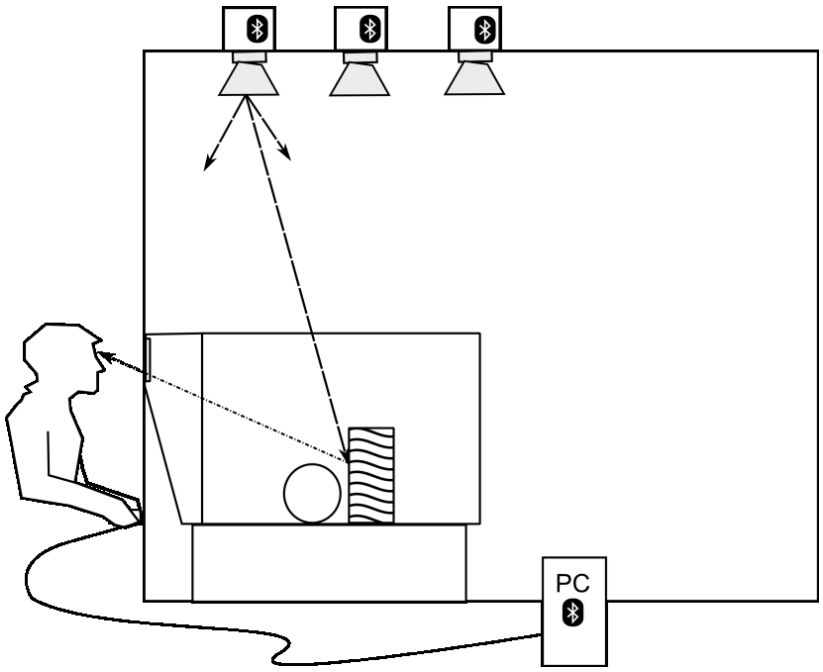

Figure 6.1. The illuminator room with luminaires, illuminating the walls of the room and viewing box.

Stimuli

Metamers from the illumination sets generated for Experiment 2, measured from the illuminator dome, were generated using the quadratic programming technique described in Chapter 2, resulting in spectra with chromaticities each < 1 $\Delta E_{u^*v^*}$ from the originals. The luminaires used had a differing number of primaries, as discussed previously in Chapter 2, which prompted the regeneration of metamers. Fifty comparison illuminations in the bluer, redder, greener and yellower directions were generated, along with a D67 target illumination.

The scene geometry of a Mondrian scene was generated using procedural generation, in python, using the Rendertoolbox package and Blender (Heasly et al., 2014). This established a list of surfaces with vertices information, which were then tagged with one of 50 material names, as can be seen by Figure 6.2.

Firstly a list of 'material' names, that would later be used as variable names for surface reflectance functions, were created (50 different Mondrian SRFs names, and a 'black' material name); a limit of 50 different materials was imposed as longer scene generation times for creating the geometric structure of the scene were experienced with larger numbers of materials. Next, 1000 planes were created with constrained coordinates, with their X positions randomised over a spread of 80cm (the viewing box back width), the Y positions randomised over a spread of 50 cm (viewing box height), and the Z positions held fixed with a random addition or subtraction of between 0 and 0.0001mms, to prevent two bodies occupying the same physical space, but with so little difference that they were considered flat; this procedure created a Mondrian back wall. The same

procedure then produced a Mondrian floor by generating 1000 more tiles and rotating the pitch by 90 º to make the plane flat to the viewing box floor. The viewing box was then created surrounding the Mondrian tiles, with a depth of 22cm (see Figure 6.2 for a schematic of the wire mesh), and a hole of $18cm^2$ was cut in the centre of the front face to allow a view inside of the box. The camera was then placed 76cm away from the centre of the box, at a height of 10cm and offset by ±3.2 cm for rendering left and right eye views. Each of the Mondrian planes was then assigned a material name from the list at random, and the box was assigned the black material name. Finally, four cubes, elevated at the top of the box 15cm x 10cm in size, were created and assigned as area lamps, with their spectral power distribution set to the variable name *illumination*. A note of the random seed was made such that it could be created again if needed; further objects could then be created each with a sculpted or random seed (so scenes could be merged together).

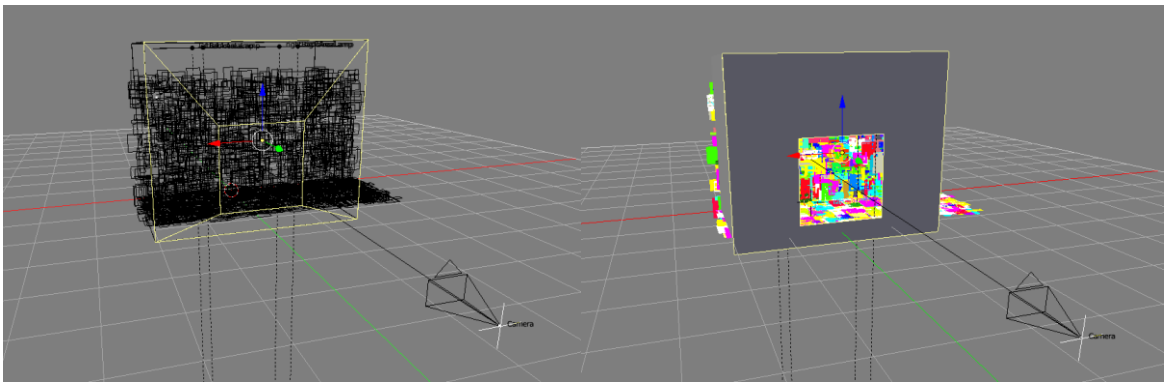

Figure 6.2. Screenshots of the procedurally generated objects in Blender: Left, wireframe planes (without materials visible); right, the scene with materials applied (colour for identification only, not actual renders). Dotted lines downward indicate the position of the light sources, and the camera is visible.

Once the scene geometry had been created the source files for Rendertoolbox were prepared such that the variables assigned in the scene file were populated randomly with non-grey, Macbeth colour checker chart values and illumination spectra. The scenes were then passed, with applicable variables, to the Mitsuba rendering programme which produced raytraced images for the scene, for the two camera vantage points. The resulting output was a 1920x1200x39 dimensional array where at each pixel location l_{xy} was the reflected spectra $l_{xy}(\lambda)$ with power values for that pixel from 400nm to 780nm at 10nm resolution. This hyperspectral image was then converted to a regular image 1920x1200x3, by calculating the tristimulus value of that point, using the CIE 1931 colour matching functions, and further converting those chromaticities

to device dependent RGB; an example of the rendered output can be seen in Figure 6.3, compared to the real Mondrian scene. The maximum size of patches in visual angle at the viewing distance of 98cm (back wall of box) was ~6 º of visual angle.

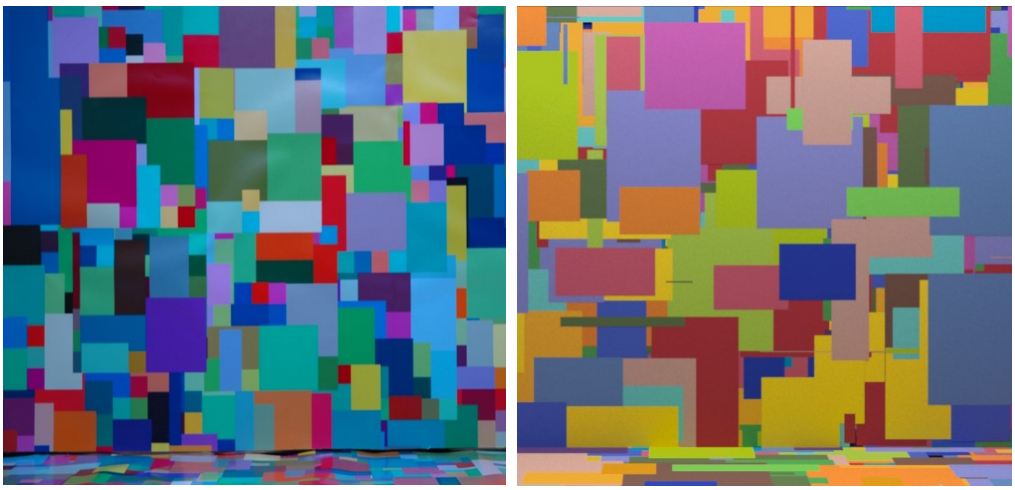

Figure 6.3. Left: Real variegated scene, under D67 illumination (RGB camera photograph). Right: Rendered variegated scene under simulated D67 illumination (converted to device dependent RGB). Rendered patches use Macbeth Colour Checker and real variegated scene surface reflectance functions. Comparative differences in scenes are due to image capture procedures.

Procedure

Participants in both conditions were given the same standardised instructions, identical to those in Appendix 4, used in Experiment 2.2. Each participant was asked to match which of two illuminations, presented consecutively, was closest to the target illumination. Participants completing the real-scene condition performed the exact procedure outlined in Experiment

2.2, which will not be repeated here. Participants at the University of Pennsylvania, who completed the simulated scene condition, were seated before the stereoscopic rig. They observed the 3D Mondrian scene, under a simulated D67 illumination through the viewing holes; this was the target scene. In each trial, the participant viewed the target scene for 2000ms, followed by a 400ms dark period where a black screen was shown to simulate the Mondrian without illumination. Following this dark period, a Mondrian scene was shown under another simulated illumination for 500ms, which was either the scene under the target illumination or under an illumination up to $\Delta E_{u^*v^*}$ away, as determined by the staircase procedure. A further 400ms dark period followed before the second comparison scene was presented, again under the target illumination if the illumination change had previously been presented or that comparison scene if the target had previously been presented. The position of the comparison, with shifted illumination, was chosen randomly by the computer to be in comparison slot 1 or 2 on each trial. Once the second comparison had been shown, a final dark period for that trial was presented, to allow the observer to give their response. The participant responded by pressing one or two marked on the gaming pad, at which time the next trial began; trials were self-paced. Participants completed trials until each of the three staircases per chromatic change direction had completed, to a maximum of 600 trials, or 6 reversals on each staircase.

Thresholds

Thresholds were obtained by averaging the last 5 reversals of each staircase, with mean threshold per chromatic direction determined by averaging the thresholds from each of the three staircases per direction, in the same way as documented in Experiment 2.2.

Results

In each condition bluer illumination thresholds were the highest (real μ = 13.99, sim. μ = 15.79) with red illumination thresholds lowest in the simulated scene (μ = 6.66) and green illumination thresholds in the real scene yielding lowest thresholds (μ = 6.63), as can be seen in Figure 6.4. Overall there was no significant difference between real and simulated scene conditions $F(1,19) = .440$, $p > .05$. There was a significant difference between illumination directions across conditions $F(3,57) = 28.510$, $p < .001$; Tukey's test revealed that red and yellow illuminations were not significantly different from each other and that blue and green were significantly different at the .05 level. There was a significant interaction effect between condition and illumination change direction $F(3,37) = 3.507$, $p < .05$. This difference is highlighted by the differences in red and green thresholds between conditions, which are significant at the .05 level, as revealed by the LSD test.

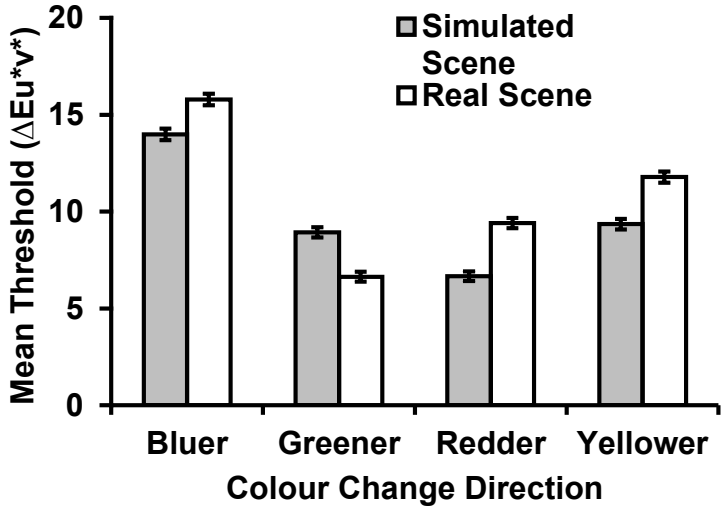

Figure 6.4. Mean observer thresholds for the two scene conditions, from the last 5 reversals of each staircase. Error bars show ±1 SEM (n= 21).

Discussion

Firstly, it is clear that participants are able to perform the task in the simulated scene, and these data suggest that observers are performing the same task in both conditions. It also appears that these thresholds are an accurate assessment of general discrimination ability by the human visual system, as similar thresholds are obtained for both sets of participants on average, whether in Europe or the United States.

It is clear that the observed pattern of thresholds in the real scene condition can be achieved with only the information from the product of the surface reflectance functions and the illumination on the scene. Therefore, the hypothesis that the observed thresholds could be due to some extraneous variable in the real scene condition can be rejected. Moreover, information such as gloss (Granzier et al., 2014; Xiao et al., 2012; Yang & Maloney, 2001), and dynamic range appear not to affect the established bias in thresholds.

These data show a systematic higher threshold for illumination discrimination of bluer illuminations as previously observed; on the other hand, the balance between redder and greener thresholds appears almost reversed. This reversal in the pattern of the lowest thresholds may be coincidental, but it may also be due to a difference in the distribution of scene surface reflectance functions which arose as a result of the surface sampling process. In fact, the simulated scene appears to have a higher frequency of yellow and red patches in the scene, compared to the real scene distribution (see Figure 6.3). An explicit comparison of the two scene chromaticities demonstrates that there is a shift in average scene chromaticity. Figure 6.5 plots the average scene chromaticities under each of the illuminations, for both the real and simulated scenes. The real scene averages were calculated from the hyperspectral image data, the simulated scene from the hyperspectral data produced by the renderer. In both cases the spectra at each pixel were multiplied by the colour matching functions (Stockman & Sharpe, 2006), and then the mean taken for each channel. While the scene average does not explain the asymmetric bias in thresholds, as previously established (Pearce et al., 2014), it may shift absolute thresholds in a similar way as grey surfaces can reduce them. The role of scene averages should be examined in greater detail now that the scene can be tightly controlled on a computer display. One prediction based on the evidence above, suggests that illumination discrimination should be poorer around the adaptation point, contrary to the evidence that chromatic discrimination is increased around the adaptation point (Krauskopf & Gegenfurtner, 1992).

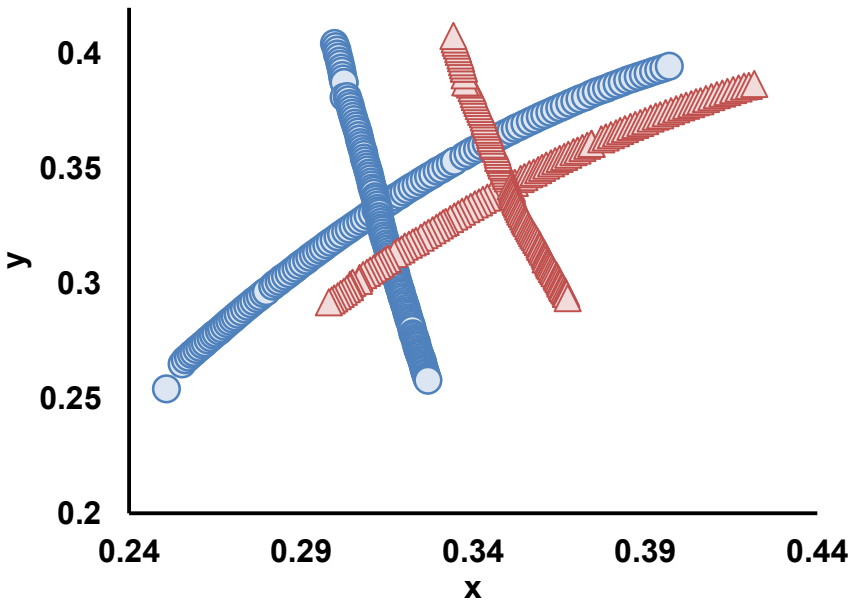

Figure 6.5. Plot of mean chromaticity of hyperspectral image of real Mondrian box under each of the illuminations (blue), and mean chromaticity of the hyperspectral render of simulated Mondrian box under rendered illuminations (red). Plot in CIE 1931 Yxy space.

One important observation is that the participant doesn't need the actual spectra and surface reflectance functions of the surfaces and the illumination information to perform the task. The inferred achromatic surface hypothesis (Brainard et al., 2006) assumes that the observer has an internal representation of these functions, that is inferred. These data support this hypothesis in so far as, if indeed the observer's visual system does have an understanding of surface reflectance functions and illuminations (at some level), then indeed it is inferred; moreover, the inferred achromatic surface hypothesis predicts the observer inference is the same for both scenes, due to comparable performance. As the simulated scene is generated using a physical model, the relationship between the surface chromaticities is preserved, suggesting that only the cone-excitation ratios are required to perform illumination discrimination, a prediction made by

relational colour constancy (Nascimento & Foster, 2000), as this is the only stable characteristic in both scenes.

It remains elusive why a bias is present for illumination discrimination thresholds and not for chromatic discrimination thresholds for surface colour changes. Under normal circumstances, illumination changes are generally global changes, spanning multiple surfaces, and chromatic changes due to surface reflectance changes are largely local contrast changes (Craven & Foster, 1992). An object, changing colour while the rest of the scene remains the same, should yield discrimination thresholds as seen in chromatic discrimination experiments if the brain interprets the change as a surface colour change; this would be evident with no bias for a particular colour direction within the thresholds. To attempt to establish surface discrimination thresholds using the same paradigm, a sphere was added to Mondrian scene geometry. Observers were asked to perform the same task as before, but on the sphere changing colour rather than the illumination, with the rest of the box remaining constant under D67 illumination; the sphere's colour change on each trial was the same as under the equivalent illumination change in the illumination discrimination task, when the colours of all of the surfaces changed under the global illumination change. It was predicted that if the bias in the illumination discrimination task arises from the difference in mechanisms used for global illumination versus local surface measurements then the bias would not be observed in the sphere surface colour discrimination task.

Experiment 4.2

Methods

Ethics

This experiment was conducted at the University of Pennsylvania, designed by Bradley Pearce, Ana Radonjic and David Brainard, programmed by Bradley Pearce and conducted by Ana Radonjic and Hilary Dubin, in accordance with APA ethical principles. Written consent was obtained as previously.

Participants

9 naïve participants (ages 18-24yrs, mean = 19.8, 7 female), who had not completed any of the experiments previously, were recruited through the University of Pennsylvania's research participation program for course credit. Participants were tested for normal colour vision with the Ishihara colour plates, and had normal visual acuity, one participant (F, Age 19) wore corrective eyewear.

Design

A two-alternative forced-choice, repeated-measures, within-subjects design was utilised; using a 1-up, 2-down, adaptive, interleaved staircase procedure; the same as in the previous experiment. The independent variable was the chromaticity of a comparison sphere suspended in the Mondrian viewing box, which was between 0 and 50 $\Delta E_{u^*v^*}$ from the target sphere. The dependent variable was the observer's threshold for discrimination, at which they selected the 0 $\Delta E_{u^*v^*}$ comparison as opposed to the non-zero $\Delta E_{u^*v^*}$ comparisons as the matching scene, above chance.

Apparatus and Stimuli

The same stereoscopic viewing apparatus, experimental code and hardware were used as in the previous experiment (see experiment 4.1). The same Mondrian geometry was used as the previous experiment; however, a sphere was added with ~5 º of visual angle in face diameter. Nine scenes of the Mondrian box were then rendered for each of the left and right views of the box, in which the surface reflectance of the sphere was specified so that the sphere's chromaticity, under D67 illumination equalled either (1) that of the monitor's R, G and B channels each set to max power; (2) that of the R, G and B channels set to 10% power, and (3) R, G and B channels at one percent power. Then, a linear combination of plus or minus any part of those images could be taken to achieve a scene where the sphere had a specified chromaticity, and the rest of the scene remained under the same illumination; called partitive mixing as described by (Xiao et al., 2012). An example of scenes where the chromaticity of the sphere is set to have that of a surface with 40% reflectance between 400nm and 780nm, under D67 (grey), can be seen in Figure 6.6. As is visible, the surrounding scene luminance is held constant, and the luminance of the sphere was held constant as the chromaticity changed.

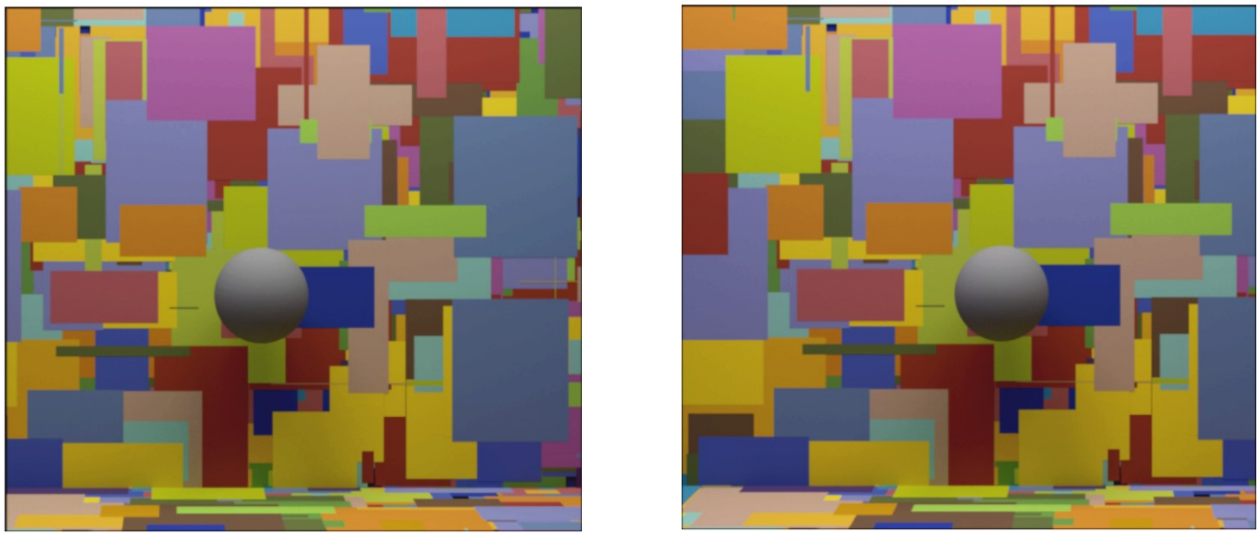

Figure 6.6. Partitive mix images of the left and right stereoscopic view of the Mondrian viewing box, where the scene is as it would appear under D67 illumination and the sphere is set to have a reflectance of 40% across the visible spectrum (the sphere and surround are both under D67 illumination).

Procedure

Participants were given the same instructions as before, however participants were asked to choose which one of the two scenes was the closest match to the target scene; participants were explicitly asked to focus on the sphere within the centre of the scene.

In each trial the participant saw a target scene, where the central sphere was neutral, and reflected the D67 illumination chromaticity, for 2000ms as before. Then a dark period of 400ms preceded the first comparison scene, which was displayed, as before, for 500ms. A second dark period of 400ms then followed before the second comparison scene was presented for the same time as

the first. A final dark period followed where participants gave their response, 1 or 2, on the gamepad, which prompted the next trial. One of the comparisons was always the same as the target, the other had a chromaticity between 0 and 50 $\Delta E_{u^*v^*}$ away from the target sphere, which varied along the daylight locus (bluer or yellower) or redder and greener along the orthogonal locus. The $\Delta E_{u^*v^*}$ contrast was chosen by the adaptive staircase procedure. Participants, as before, completed 600 self-paced trials, or 6 reversals on each staircase, whichever was lower. Participant thresholds were established as before.

Results

There was an observed asymmetry between thresholds, with blue thresholds highest overall and green thresholds lowest (see Figure 6.7); thresholds for each colour change direction were significantly different from each other $F(3,48) = 25.993$, $p < .001$. Post-hoc pairwise comparisons revealed that all threshold comparisons were significantly different at $p < .05$ level, except yellow vs red thresholds ($p = .643$). There was no significant difference between the surface discrimination thresholds and the illumination thresholds for the previous monitor experiment (Experiment 4.2) $F(1,16) = 2.299$, $p = .149$.

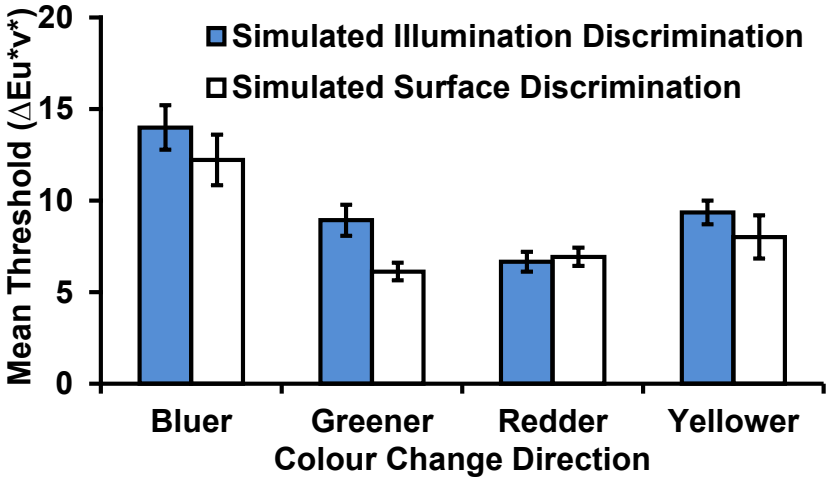

Figure 6.7. Mean observer thresholds for the two simulated conditions, computed from the last 5 reversals of each staircase. Error bars show ±1 SEM (N = 18, n = 9), illumination and surface discrimination data from two different samples as described above.

Discussion

Participants were able to perform the task as before; however, these data do not support the prediction that thresholds should differ from those obtained in the global illumination discrimination task. The results obtained from the two tasks -- global illumination discrimination and local surface discrimination -- are not statistically different; thus, these data do not conclusively support a distinct difference in task. Therefore, it is a matter for discussion whether either or both tasks are indeed illumination discrimination tasks.

These data do suggest that in both experiments the observer has the same adaptation point; and despite the change in field size of the target, and neutral

point, both known to affect colour constancy (Werner, 2003), thresholds remained the same.

One possibility is that the sphere could be determined to have undergone an illumination change, as it appears achromatic at the adaptation point and is segregated in depth. Hurlbert and Wolf (2004) discuss how colour contrast between the surround and a test patch might indicate that an object is under a different illumination.. . If this is not the case, and the task is in some way a surface discrimination task, it remains to be explained from where the asymmetry in discrimination thresholds within chromatic axes arises. Making the target surface a flat, matte patch embedded within the scene would remove depth cues, with the possible side-effect of chromatic induction from the surrounding scene; it would be predicted for this task, on the basis of chromatic discrimination thresholds from the same adaptation point, that no asymmetry would be present.

Chapter 7:
Immersive Colour Constancy

While many simultaneous colour constancy experiments have investigated colour appearance using two scenes under separate illuminations (Smithson, 2005), the observer typically views the two scenes from an external vantage point, which is itself in a third illumination framework (typically with very dim or nearly non-existent illumination); adaptation is therefore likely to be a combination of illuminations. In other experiments, the observer views a single scene under a single illumination, for example to make achromatic adjustments or assessments (D'Zmura, Rinner, & Gegenfurtner, 2000; Kraft & Brainard, 1999), but again is situated outside the illuminated chamber. It therefore remains a question as to whether immersion in the "adapting" illumination affects the state of adaptation. It also remains unclear whether an observer immersed in the illumination rather than viewing the scene externally will have the same level of constancy for surfaces under changing illuminations, for the two conditions .

Data from the surface discrimination experiment using the grey sphere suggests that if that task is indeed governed by the same adaptational mechanism as the illumination discrimination task then this adaptational mechanism is relatively insensitive to the size of the adapting field, since the same result is obtained for a chromaticity change across the isolated ~5-degree sphere and the entire ~44-degree Mondrian wall. This result seems counter to those from chromatic discrimination experiments in which changes in field size alter the results, e.g. in the determination of colour matching functions (Wyszecki & Stiles, 1982) or the obtaining of constancy indices under certain conditions (Murray, Daugirdiene, Vaitkevicius, Kulikowski, & Stanikunas,

2006). On the other hand, it is premature to assume that chromatic adaptation alone governs the illumination or surface discrimination task. Nonetheless, we may ask whether changes in the adaptational state, such as would result from complete immersion in the "adapting" illumination, affect the thresholds in the illumination discrimination task. If adaptational mechanisms govern performance on the illumination discrimination task, then one might expect constancy, and indeed illumination discrimination thresholds to change when the observer is immersed within the illumination as opposed to viewing the scene through a porthole. To test this hypothesis, participants completed the same illumination discrimination task while viewing the Mondrian scene as before; however, they completed the task inside the lightroom, and were illuminated from above along with the scene. It was predicted that if full field adaptation affected adaptation state, the consequent change in adaptational state would be reflected in illumination discrimination thresholds, with better constancy reflected by higher thresholds.

Experiment 5.1

Methods

Ethics

This work has been reported in Hurlbert, Pearce, Mackiewicz and Finlayson (2014). This experiment was conducted according to APA ethical principles, and was granted ethical approval from Newcastle University's FMS Ethics

Committee (reference number 00312). All participants gave written consent after reading standardised instructions.

Participants

Six naïve observers (3 female; aged between 20-28 yrs) participated in the experiment. Participants were post-graduate researchers from within the Institute of Neuroscience and participated as part of research projects. All participants had normal or corrected to normal colour vision.

Design

A two-alternative forced choice, repeated measures, within-subjects design was employed utilising an interleaved, 1-up, 2-down staircase design as previously. The independent variable, as before, was the perceptual distance of a comparison illumination from a target illumination in the CIELUV colour space; the dependent variable was the observer's measured threshold for illumination change, as determined by the last 5 reversals of each 1–up, 2–down staircase. Three interleaved staircases were used for each of the four illumination change directions: bluer, redder, greener and yellower. Each staircase had a maximum of 6 reversals.

Apparatus and Stimuli

The Mondrian box, as described previously – with dimensions (71cm x 77cm x 47cm), with front removed, was placed inside a $2m^3$ room (lightroom) with white-painted walls (mean Yxy: x=0.31 y=0.32 under D67); the box constituted ~26° of viewing angle, of a full field view of the back of the lightroom, itself ~68° of viewing angle.

Three prototype tuneable LED luminaries were positioned in the ceiling such that they diffusely illuminated the room and its contents. Black polymer, faux-fur was used to cover a small stool and stand for the Mondrian box which had approximately uniform reflectance of ~0.8%. A blacked-out XBOX 360 controller was used for participants to provide feedback to a computer which stood outside of the room, running Windows 7 and custom software written in MATLAB to control the tuneable LED luminaries and experimental code. The computer also provided auditory feedback to the observers via black headphones.

The experimental illuminations were the same as used in Chapter 4; ranging from ±50 $\Delta E_{u^*v^*}$ from D67 along the daylight locus (blue and yellow), and along the locus of correlated colour-temperature 6700K (red and green). All illuminations were isoluminant with a CIE Y of 250 cd/m^2 (as measured from a white calibration tile positioned, placed in the centre of the viewing box ~2m away from observer's seated position, and ~1m away from the luminaires' point of maximal brightness) and maximum deviation of ± .05 cd/m^2. All illuminations were regenerated from basis functions taken using a CS-2000 spectroradiometer. A schematic of the setup can be seen in Figure 7.1, along with a photograph of the scene.

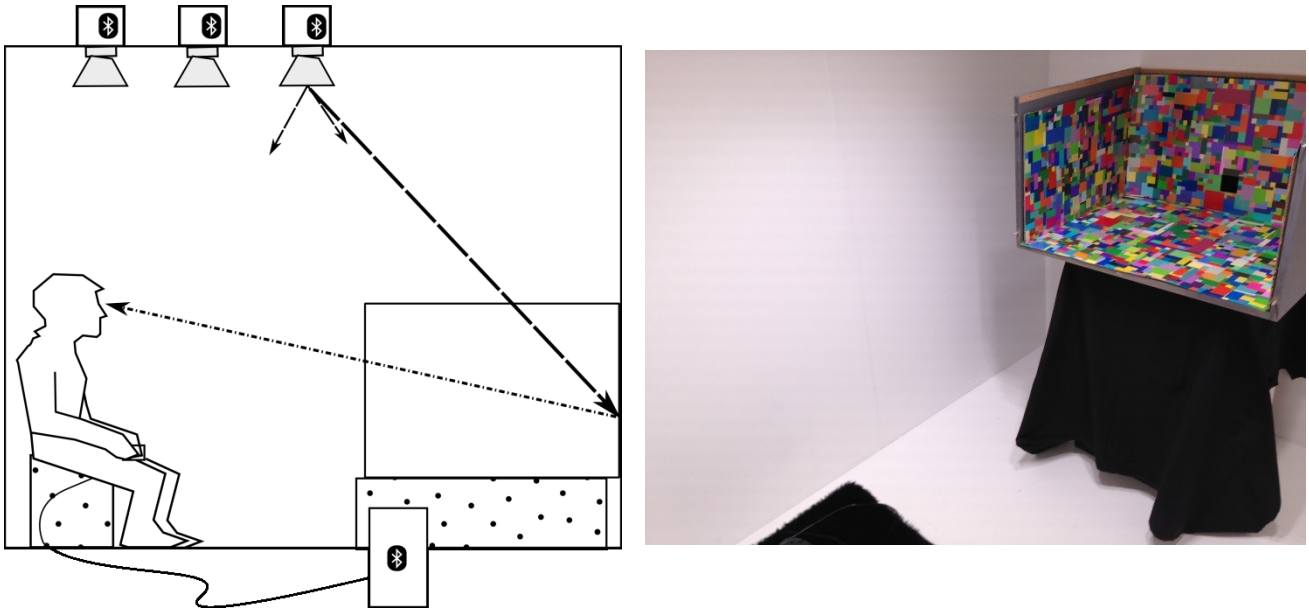

Figure 7.1. A schematic of the viewing setup and apparatus (left), along with a photograph of the Mondrian box situated within the lightroom.

Procedure

The illumination inside the chamber was set to D67, the target illumination for all trials. Participants were seated on the stool at a distance of ~150cm away from the front of the viewing box, inside the lightroom, and asked to look inside the viewing box. Participants were given the same standardised instructions as before, and were instructed that when the experiment started, and between trials, the illumination in the room would be turned off. The lights were then turned off and the experiment began. The trial format was as before, with a target illumination, D67, turned on such that it illuminated the lightroom, viewing box, and the observer for 2000ms. A dark period of 400ms separated two comparison illuminations. As in previous experiments, one comparison was the target, the other was an alternative light with a perceptual distance up to 50 $\Delta E_{u^*v^*}$ away from the target illumination. Each illumination, and the final dark

period after the comparisons were presented, was accompanied by an auditory cue; observers gave their responses as before, in darkness.

The comparison illumination for each trial was selected by the staircase based on the previous contrast and the current step size for that staircase. Each staircase was indexed and one was chosen randomly at the start of each trial. The experiment continued until a maximum number of 600 trials was reached, or each staircase had finished reversing.

Thresholds

Thresholds were determined as before, by averaging the last 5 reversals of each staircase after the nominal staircase step was converted to Euclidean distance ($\Delta E_{u^*v^*}$ units) from D67.

Results

Mean discrimination thresholds can be seen in Figure 7.2, along with the thresholds from Experiment 2.2 where observers were not immersed within the lightroom; there was no significant difference between the thresholds obtained from observers inside or outside the lightroom $F(3,14) = .270$, $p = .846$. There was a significant difference between the illumination colour change directions $F(3,14) = 14.038$, $p < .001$, as previously observed.

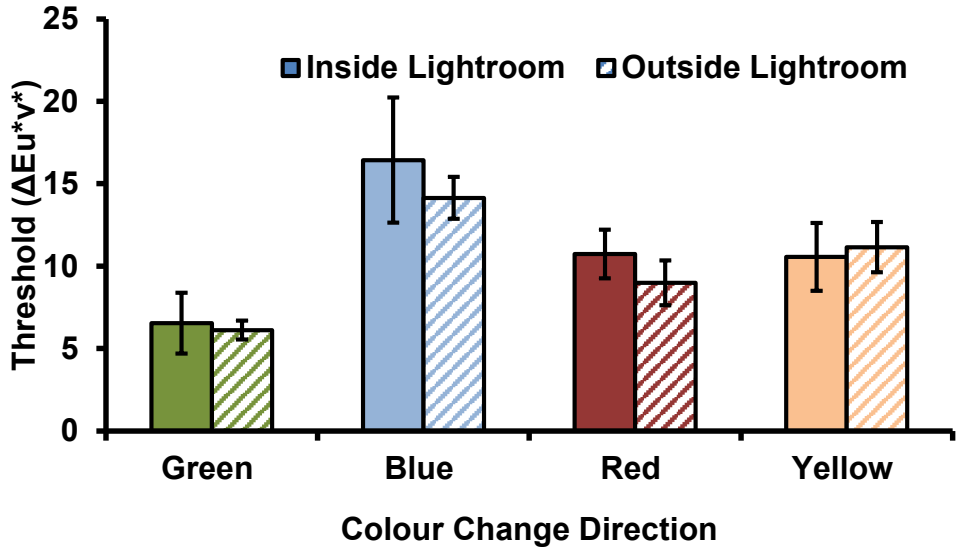

Figure 7.2. Mean illumination discrimination thresholds for observers sitting inside and outside the lightroom. Error bars show ±1 SEM (n = 6).

Discussion

As predicted, illumination discrimination thresholds were not significantly different from being inside or outside the lightroom. This suggests that the illumination discrimination thresholds are not field size sensitive, at least at the resolutions tested here. The white walls of the lightroom did not lower illumination discrimination thresholds as the achromatic objects did previously; this may be due to there being no achromatic objects embedded within the scene making judgments based on the local contrast more difficult; however on face value the threshold means and overall variance in thresholds is higher for those obtained inside the lightroom.

Equally, as demonstrated previously, the dynamic range within the scene does not affect thresholds; in the outside condition the brightest patch was the maximum flux within the scene; however, for the inside the lightroom condition,

maximum flux was never inside the viewing box because of the white walls; yet, thresholds are unmoved, contrary to the predictions of retinex (Finlayson, et al., 1997). This adds further evidence that the mechanisms that modulate this task are not predicted by cone-contrast in general, as a perfect white in the scene, such as the white walls, signal the exact illumination change; however, local contrast, such as that between the chromatic and achromatic surfaces within the scene may indeed cue the visual system, as described in Chapter 5. Also, while the preservation of cone-excitation ratios appears to be important to identifying an illumination change, there must be a limit to the amount that additional surfaces and viewing angle can cue these mechanisms (Foster, Nascimento, & Amano, 2005) as field size as described here has no effect on thresholds. It may be that only surfaces that violate cone-excitation ratios are considered when the visual system is assessing whether a surface or the illumination has changed, and the rest of the scene is not factored into the computation, a possible explanation for why the surface and illumination discrimination thresholds in these experiments appear similar.

While it is still unclear whether the task being performed by the observer is one of illumination discrimination, one way to parse the tasks would be to show, using a similar paradigm, differences in thresholds for varying field size for surface and illumination discrimination tasks.

Chapter 8:
General Discussion

Are colour constancy mechanisms biased towards particular illuminations?

The first research question posed in the introduction to this work was whether colour constancy operates better under some illuminations than others; and, specifically, whether constancy mechanisms are biased towards the illuminations under which they have evolved, following the so-called ecological hypothesis of colour vision (Cecchi, Rao, Xiao, & Kaplan, 2010; Sumner & Mollon, 2000). In the first experiment, the generated metamers were created to be smooth and as close a spectral and colorimetric match to daylight as possible with the tuneable LED technology. These constraints were also placed on the generation of novel illuminations with CCT of 6700K, the green and red locus. Under these illumination changes, an unbiased visual system should have exhibited the same illumination discrimination thresholds for each of the bluer, yellower, greener and redder illumination changes as the distances between alternatives in cone-contrast coordinates were roughly equal in each of the Red-Green and Blue-Yellow Axis. Indeed, this was not the case, and the accuracy for illumination matching was lower, and illumination discrimination thresholds, higher, (as established in Experiment 2) for bluer daylight illuminations that the non-daylight counterparts in the Red-Green colour direction.

Previous models of colour constancy, most notably max-flux and grey-world (Hurlbert, 1989, 1998), do not explain the bias observed in the illumination discrimination thresholds, as the location of maximal change was both different and equivalent in magnitude for each illumination change colour direction; and,

the scene average was different in each scene (Mondrian or Grey in Experiment 1), yet observer threshold ratios were approximately equivalent.

The argument that the visual system is biased towards a specific illumination *chromaticity*, regardless of the spectra that constitutes it, is one of poor foundation. As shown in detail by Figure 4.1, a MacAdam-like ellipse for the observer's illumination discrimination thresholds, thresholds for bluer illuminations were twice as large as any other illumination colour direction. These thresholds make a statement about how far perceptually, as defined by the CIE Lu*v* colour space, the illumination can change before the surfaces within the scene, by virtue of the illumination change, signal that change. However, an arbitrary number of metamers can give the same illumination chromaticity (Finalyson, Mackiewicz, Hurlbert, Pearce & Crichton, 2014); as can be seen in Figure 8.1, both boxes are illuminated by D67 metamers with varying spectral power distributions; the almost uniform grey box exterior appears the same in both scenes as it reflects the metamers faithfully in both instances, yet the scenes look radically different and colours are not constant between the scenes.

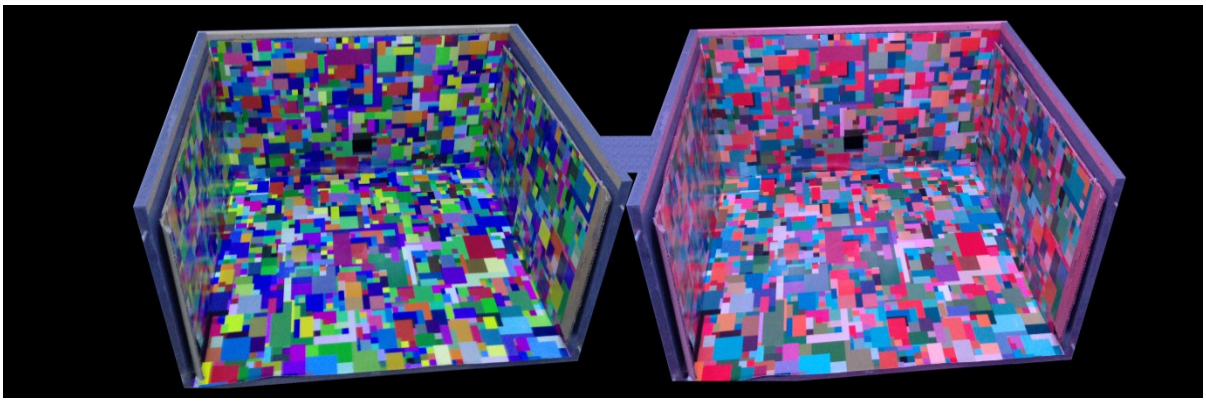

Figure 8.2. Photographs of the Mondrian box under two D67 metamers. The image has been altered by extending the image of the wooden box border between the boxes to demonstrate that the colour of the illumination is the same, but the chromatic patches within the Mondrian pattern look very different under the two illuminations.

It is clear that colour constancy cannot operate effectively under these conditions in which although the illumination is essentially the same in chromaticity, nonetheless the scene average chromaticity, max-flux and relationship between the surfaces are not. This concept is examined most closely by Logvinenko (2015), who describes a hue circle of colours from a collection of surfaces becoming metameric under two illuminations which themselves are metameric. The predictions made by Logvinenko and colleagues (Logvinenko, Funt, & Godau, 2014; Logvinenko, 2015) are demonstrated here, with the two scenes in Figure 8.2, under which an observer would be expected to achieve differing states of adaptation by virtue of the surface ensemble. This raises serious difficulties for colour constancy if the possible illuminations are left

unconstrained as there is any number of possible solutions for any given surface colour, by virtue of any infinite number of spectra for any scene average chromaticity.

However, under natural daylight conditions, the illumination is very predictable, depending on time of day and atmospheric conditions (Hernandez-Andres et al., 2001; Webster et al., 2007); therefore, the rendition of surfaces under these illuminations is predictable. That is, as daylight changes, surface colours retain their approximate cone-excitation ratios (Foster & Nascimento, 1994), and colour rendition is comparable (Sandor & Schanda, 2006). These data therefore support a mechanism of colour constancy that is biased towards daylight 'bluer' chromaticities that are produced by smooth changes in spectra which leave surface cone excitation information roughly intact. This is supported by the physical evidence of spectra and surface variations found in natural scenes (Golz & MacLeod, 2002; McDermott & Webster, 2012; Nascimento, Ferreira, & Foster, 2002) and computational models of colour constancy that assume the preservation of cone-excitation ratios between illumination changes (Foster & Nascimento, 1994). Furthermore, computational models of colour constancy that assume a probability of an illumination upon a scene, as described by Finlayson, Hordley and Hubel (2001), could describe illumination discrimination thresholds by using a weighted probability for bluer illumination changes.

In the real and simulated scenes, described in experiments within Chapter 6, it is clear that those spectra that produce small changes in surface appearance

are not directly represented by the visual system; that is, the sensor information at each pixel is represented only, as the theoretically identical spectra illuminating the two scenes are produced by artificial primaries in the simulated scene which do not match the daylight basis functions, yet, the observer thresholds are comparable. This lends weight to the conclusion that mechanisms that produce a bias for the blue illumination change thresholds reside in the cortex because the sensory data, as revealed by the hyperspectral camera for the real scenes, and the images produced by the renderer, show cone-contrast signals that are equally discriminable for each illumination change direction, suggesting that processing of the image signal in the retina and thalamus – at the lowest levels of the colour vision pathway – do not reflect this bias. The results are suggestive, but not conclusive, of the existence of a processing bias at higher levels which weights constancy mechanisms asymmetrically along different chromatic directions. The green illumination thresholds add weight to this hypothesis, as they are not significantly different from chromatic discrimination thresholds (reported in the colour matching literature and measured by the CCDT), with responses to changes not diminished relative to other chromatic directions; the low green illumination discrimination thresholds and high blue illumination thresholds in this respect are not explained by any current sensory model of colour constancy.

The differences observed between the chromatic and illumination discrimination thresholds may be mediated by spatial scale, as the illumination change is a full field contrast change, whereas the arrow's chromaticity change only required a single change at high spatial frequency to be detected. Target

patches of different spatial scale relative to an adaptive field (background) have been shown to modulate colour constancy (Hansen et al., 2007), with higher levels of constancy observed for smaller target patches on larger surround fields; perhaps explaining the difference in absolute levels of constancy between the Mondrian and grey boxes. The experimental paradigm does not allow examination of the effects of spatial contrast; however, the Mondrian scene has variations in both chromaticity and luminance as a function of spatial frequency, each having differing contrast sensitivity functions (Mullen, 1985), which are much less prevalent in the grey condition, and yet the bias between the illumination directions is preserved.

It is important to state the scope of the bias for thresholds of illumination changes in this task. The illumination discrimination task, in general, measures colour constancy immediately following an illumination change; that is, the absolute upper level of constancy at that point. This paradigm cannot make a statement about adaptation over time, but may only predict that if the illumination change has not been detected immediately, colour constancy may be considered complete moving forward temporally. Moreover, the paradigm does not allow observers to attribute change to a surface rather than the illumination, as both the scene is static and the instructions are to detect an illumination change. As has been demonstrated by Arend et al. (1991), levels of constancy were modulated by instruction. It is a legitimate question to task whether the observered asymmetry in thresholds would be present for a comparable surface judgment task which controlled contrast changed by spatial frequency.

Previous studies of colour constancy, using real scenes under changing illuminations, show reasonable constancy indices (Kraft & Brainard, 1998; Brainard, 1999); demonstrating that colour constancy can be good under arbitrary illuminations. However, as shown by an achromatic matching task, using the same Mondrian box and illuminations detailed here, constancy indices were better for daylight than for illuminations on an orthogonal locus over a 10 second period of adaptation (Crichton, Pearce, Mackiewicz, Finlayson, & Hurlbert, 2012); moreover, better colour constancy has been shown for illumination changes along the bluer-yellower direction of the daylight locus than for redder-greener shifts, albeit with a symmetric difference in each axis (Worthey, 1985).

Can familiar objects cue colour constancy mechanisms towards particular illumination changes?

The second research question raised in the introduction to this work was whether certain objects, by virtue of their familiarity, affect colour constancy mechanisms. In the first experiment it appeared that neither the fruits nor chromatically matched objects affected illumination discrimination. Indeed, if one assumes that colour constancy exists to aid object recognition and to facilitate object discrimination, then it would make sense for objects not to facilitate constancy mechanisms, as such mechanisms should occur before object recognition. However, it does appear, from the experiments conducted in Chapter 5, that surfaces that are not highly chromatic, most notably the grey block, are able to cue the visual system to an illumination change. Both the

lower thresholds and the higher discrimination accuracy for the grey box in Experiment 1 support this hypothesis.

The signal change from the grey surface was not always as great as from any particular square within the Mondrian scene for each illumination change, so this suggests that it is the reliability of the surface as a cue rather than the neutrality of the surface that indicates the illumination to the visual system. That is, for a surface to inform colour constancy mechanisms, the approximate surface reflectance function must be known, or the surface should approximately reflect the illumination for arbitrary illumination changes, as implied by the equivalent illumination hypothesis Xiao et al. (2008). Granzier and Gegenfurtner (2011) report a small but significant facilitation of identifying the illumination colour on a scene when familiar objects are visible, as opposed to when they are clothed in chromatically similar paper; however, some of those objects (most notably the chocolate bar) had significant surface area that was white. Therefore, with the knowledge that achromatic surfaces may facilitate illumination discrimination, it would be incorrect to assume that the facilitation observed by the above authors was due to the familiarity of the objects. Moreover, this is corroborated by Kanematsu and Brainard (2013) who observe no effect of familiar objects on colour constancy when the presence of achromatic objects is controlled in both scenes, and these data presented here. Of course, it is not clear from these data, at what point in the processing by the visual system such surfaces are identified as being neutral in surface reflectance. This topic is not discussed here, but is a subject of legitimate inquiry.

It is expected that, over time, the signal from a scene containing only grey surfaces will be adapted more closely to an illumination, at each point over time, as the sensors have access to the illumination chromaticity. Equally, as surfaces are added to the scene, the scene complexity increases and the scene average reflectance moves more towards grey, providing the reflectances of surfaces are drawn with equal probability from a flat distribution centred on neutral; this complexity -- referred to as scene articulation (Linnell & Foster, 2002) - then provides a more reliable cue to the illumination as more surfaces are added. As documented by Linnell and Foster (2002), discerning an illumination change over a surface change is dependent on the scene average, in turn dependent on the surface ensemble. The evidence from illumination discrimination, presented here, is consistent with scene average affecting thresholds, as is observed with the simulated scene experiments; furthermore, illumination discrimination was much better when grey surfaces were present within the scene. Contrary to the predictions made by Linnell and Foster (2002), illumination discrimination was poorest with Mondrian scenes, even when the scene average was closely matched to the grey viewing box. While grey surfaces affect illumination discrimination, it is still unclear, as described by Foster (2011), how neutral surfaces affect colour constancy mechanisms over time; as such, constancy experiments containing grey surfaces should be looked at with scrutiny, as they may provide a particularly uncontrolled cue.

The effects of familiar objects on illumination and colour appearance in general will become clearer when the cortical underpinnings of colour perception are better understood. The most concrete conclusion that can be made with

regards to the effects of particular surfaces on colour constancy, is that reflectance is the most important characteristic, in combination with scene context (Hurlbert & Ling, 2006; Hurlbert & Wolf, 2003). The evidence presented here, showed similar changes in illumination discrimination for a Caucasian fake hand, and matching painted block and grey painted block; demonstrating that our ability to detect changes in illumination is dependent on the surface ensemble rather than other cognitive properties associated with those surfaces, as perceived by the observer.

What information is needed for illumination discrimination?

The real scenes used in the real-world experiments differ from the simulated scene in several ways, in addition to those already discussed. For example, the real scene contains some small amounts of gloss from the satin-like paper; the presence of gloss has been demonstrated to influence colour constancy (Xiao & Brainard, 2008; Yang & Maloney, 2001). The real scenes also contain a much higher range of chromaticities and luminances than in the simulated scene used in Chapter 6. Such cues were eliminated when the simulated scene was created. Yet a comparable threshold for real and simulated scenes demonstrates that all the available information to perform the illumination discrimination task is present in the simulated scene, which is comprised of matte surfaces and whose image irradiance signals only the information comprised from the combination of illumination and reflectance functions. This also shows empirically that the illumination perceived by the observer may be signalled by

relatively low-level information calculated from the sensor signals only. If the observer does have an internal representation of the illumination, or indeed a subjective representation, or inferred achromatic surface (Brainard et al., 2006) then it is most certainly derived at a higher level, from low-level image statistics.

Tuneable LED luminaries allow the production of an almost infinite array of metameric lights for each chromaticity (Finlayson et al., 2014; Mackiewicz, Crichton, et al., 2012). The threshold for such discrimination from one metamer to another will be determined by norm of the two metamers with respect to the surface ensemble reflecting them. It is indeed true that if there are no surfaces in the scene to reflect such changes, then colour constancy will be perfect and a threshold for illumination change infinite.

One stark conclusion on the basis of the above evidence is that colour spaces are not appropriate for determining thresholds for metamers with the same chromaticity, as there will be a threshold but the perceptual distance between the illuminations will be 0. More profoundly, spectra that are made to closely mimic the spectral form of natural daylights yield the same thresholds for three different luminaire systems and LCD display technologies, despite differences in spectral form.

What is the scope of these data?

Previous work on colour constancy has focussed specifically on subjective surface colour perception, under a small number of illumination changes (Smithson, 2005; Foster, 2011). These studies have not specifically focused on

selecting specific illumination changes, usually selecting arbitrary illuminations, with some exceptions (Kanematsu & Brainard, 2013; Delahunt & Brainard, 2004; Brainard, 1998). This work has investigated colour constancy through systematically varying the illumination over a broad range of broadband, common illuminations, for the first time achieving an objective measure of colour constancy. These threshold data show two very important findings which the current literature, inclusive of this work, cannot fully describe; thresholds for illumination discrimination are much greater than predicted by surface matching tasks.

Further illumination discrimination thresholds need to be established further along the daylight locus, with extreme blue and extreme yellow illuminations as adaptation points. This can be equally achieved by changing the scene average in simulated scenes; with these , less extreme changes have shown that illumination discrimination does depend on scene average chromaticity, as determined by surface ensemble (Krieger et al., 2014).

Traditional achromatic matches (Brainard, 1998) using the same Mondrian box and illuminations, as presented here, have revealed constancy indices which are higher for daylight illuminations than novel illumination changes (Crichton, Pearce, et al., 2012; Mackiewicz, Pearce, Crichton, Finlayson, & Hurlbert, 2012). These data were also not predicted by scene average chromaticity.

In each of the experiments documented here, successive trials used the same variegated scene. Arend and Reeves (1986) documented how simultaneous colour matches changed when observers were asked to make matches, as if the

test patch were cut from the same piece of paper as a reference patch. Because the scene reflectance was never changed in the presented studies, and the observer always understood that the change was an illumination change, a representation of patch reflectance could be inferred over successive trials. Such inferences have been suggested by similar studies by Smithson and Zaidi (2004). Therefore, absolute illumination thresholds might only be obtained once observers have inferred the scene reflectance; indeed, a hypothesis that could also be examined by perturbing the scene reflectance between trials.

These data cannot elucidate what colour names would be attributed to patches in the scene upon illumination changes; more precisely, the design only probes if the illumination change was detected. Hansen, Walter and Gegenfurtner (2007) demonstrate small changes in colour name boundaries for illuminations varied in cone contrast along blue-yellow (daylight), red, green, purple and turquoise colour directions. It is unknown whether a change in illumination might be undetected, but might still affect colour names. If illumination discrimination is indeed an objective measure of colour constancy, colour category boundaries should not shift for undetected illumination changes; more specifically, for illumination changes below threshold to an adaptation point.

Many studies have demonstrated high levels of colour constancy (Hansen, Giesel, & Gegenfurtner, 2008; Kraft & Brainard, 1999; Ling & Hurlbert, 2008b). Illumination discrimination using tuneable LED light sources allows the effects of specific surfaces, scene statistics and illumination spectral content on colour

constancy to be measured. When the illumination is monochromatic, colour constancy mechanisms cannot operate due to only luminance at one wavelength being available to the visual system from surfaces in the scene. In contrast, broadband illuminations contain polychromatic information; the surface colour information available to the visual system is perfect for a spectrum that is of equal energy at each wavelength. An arbitrary measure of colour constancy could assume a point at which the mechanism breaks from perfect colour constancy moving away from the perfectly flat illumination; see Figure 8.2, where that equal energy illumination chromaticity (Illumination E) is plotted along with the chromaticity of monochromatic light at 548nm. This work has shown that for natural illumination changes, that point is well described in colour space; however, with the advance of tuneable LED luminaire technology, and an infinite array of metameric illuminations of which to choose as a test illumination, a spectral measure of discrimination will be required to effectively characterise colour constancy.

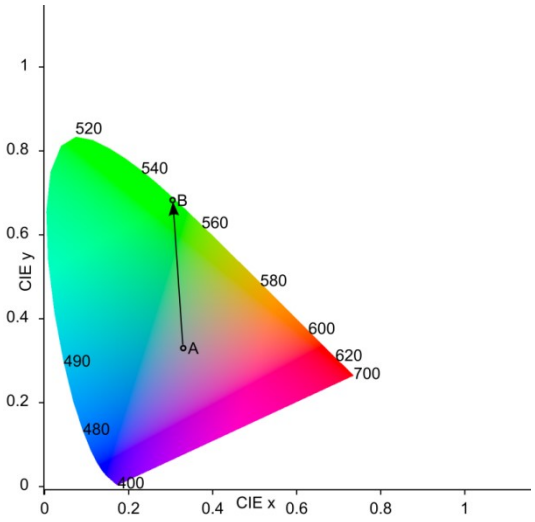

Figure 8.2. CIE 1931 chromaticity diagram with the chromaticity of illumination E (A) and monochromatic light at 548nm (B) marked. Colour constancy can operate optimally at point A, and does not operate at point B; at some point along the line of possible illumination changes between A and B, marked, constancy mechanisms will fail.

To conclude, this work stresses the importance of the surfaces within the scene, and the composition of the illumination and the composition of the illumination change. It is clear that colour constancy mechanisms are imperfect; however, it is also clear those mechanisms are sculpted by the ecological conditions under which we have evolved. These data support the hypotheses that: colour constancy is better for daylight illuminations; that achromatic surfaces are able to assist the visual system in illumination discrimination; that the familiarity of objects does not aid colour constancy; and, that the only information necessary for the visual system to determine an illumination change can be derived from light reflected from a world of patches with lambertian reflectance, whether real or simulated.

References

Amano, K., Foster, D. H., & Nascimento, S. M. C. (2006). Color constancy in natural scenes with and without an explicit illuminant cue. *Visual Neuroscience*, *23*(3-4), 351–359.

Arend, L. E., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous Color Constancy - Papers with Diverse Munsell Values. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *8*(4), 661–672. Retrieved from <Go to ISI>://A1991FE97100009

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *J Opt Soc Am A*, *3*(10), 1743–1751. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3772637

Bala, R. (2003). Inverse problems in color device characterization. In *SPIE* (Vol. 5016, pp. 185–195). http://doi.org/10.1117/12.488617

Barnard, K., Cardei, V., & Funt, B. (2002). A comparison of computational color constancy algorithms - Part I: Methodology and experiments with synthesized data. *Ieee Transactions on Image Processing*, *11*(9), 972–984. http://doi.org/Doi 10.1109/Tip.2002.802531

Baumgartner, H., Vaskuri, A., Kärhä, P., & Ikonen, E. (2014). A temperature controller for high power light emitting diodes based on resistive heating and liquid cooling. *Applied Thermal Engineering*, *71*, 317–323. http://doi.org/10.1016/j.applthermaleng.2014.06.050

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*(6764), 877–879. Retrieved from <Go to ISI>://000084482000033

Bosten, J., & MacLeod, D. I. (2012). Color confusion ellipses from absolute judgements. *J Vis*, *12*(9), 108.

Brainard, D. H. (1995). Colorimetry. In M. Bass (Ed.), *OSA Handbook of Optics: Fundamentals, Techniques, and Design* (Vol. 1, pp. 26.1 – 26.54). New York: McGraw-Hill, Inc.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis*, *10*(4), 433–436. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176952

Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *15*(2), 307–325. Retrieved from <Go to ISI>://000071669700003

Brainard, D. H., Longere, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B. (2006). Bayesian model of human color constancy. *J Vis*, *6*(11), 1267–1281. http://doi.org/Doi 10.1167/6.11.10

Brainard, D. H., & Maloney, L. T. (2011). Surface color perception and equivalent illumination models. *Journal of Vision*, *11*. http://doi.org/10.1167/11.5.1

Brainard, D. H., & Wandell, B. A. (1992). Asymmetric Color Matching - How Color Appearance Depends on the Illuminant. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *9*(9), 1433–1448. Retrieved from <Go to ISI>://A1992JK70100001

Cao, D., Zele, A. J., Smith, V. C., & Pokorny, J. (2008). S-cone discrimination for stimuli with spatial and temporal chromatic contrast. *Visual Neuroscience*, *25*, 349–354. http://doi.org/10.1167/6.13.6

Cecchi, G. A., Rao, A. R., Xiao, Y. P., & Kaplan, E. (2010). Statistics of natural scenes and cortical color processing. *J Vis*, *10*(11). http://doi.org/Artn 21 Doi 10.1167/10.11.21

Chichilnisky, E. J., & Wandell, B. A. (1995). Chromatic Adaptation Affects Achromatic Increment and Decrement Settings Differently. *Investigative Ophthalmology & Visual Science*, *36*(4), S209–S209. Retrieved from <Go to ISI>://A1995QM91500958

Conway, B. R. (2009). Color vision, cones, and color-coding in the cortex. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *15*, 274–290. http://doi.org/10.1177/1073858408331369

Conway, B. R. (2013). Color signals through dorsal and ventral visual pathways. *Visual Neuroscience*, 1–13. http://doi.org/10.1017/S0952523813000382

Cranwell, M. B., Pearce, B., Loveridge, C., & Hurlbert, A. (2014). Performance on the Farnsworth-Munsell 100-Hue Test is significantly related to non-verbal IQ. *IOVS*, *in review*.

Craven, B. J., & Foster, D. H. (1992). An operational approach to color constancy. *Vision Res*, *32*(7), 1359–1366. Retrieved from <Go to ISI>://A1992JC08800016

Crichton, S., Pearce, B., Mackiewicz, M., Finlayson, G., & Hurlbert, A. C. (2012). The illumination correction bias of the human visual system. *ARVO*, *12*(9), 64.

Crichton, S., Pichat, J., Mackiewicz, M., Tian, G., & Hurlbert, A. C. (2012). Skin chromaticity gamuts for illumination recovery. *Conference on Colour in Graphics, Imaging, and Vision*, *2012*(1), 266–271.

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. *The Journal of Comparative Neurology*, *312*(4), 610–624. http://doi.org/10.1002/cne.903120411

D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America. A, Optics and Image Science*, *3*(10), 1662–1672. http://doi.org/10.1364/JOSAA.3.001662

D'Zmura, M., Rinner, O., & Gegenfurtner, K. R. (2000). The colors seen behind transparent filters. *Perception*, *29*(8), 911–926. Retrieved from <Go to ISI>://000165403300004

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), 582–588. http://doi.org/10.1073/pnas.93.2.582

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), 731–735. http://doi.org/10.1038/367731a0

Dannemiller, J. L. (1993). Rank Orderings of Photoreceptor Photon Catches from Natural Objects Are Nearly Illuminant-Invariant. *Vision Res*, *33*(1), 131–140. Retrieved from <Go to ISI>://A1993KE83600015

Delahunt, P. B., & Brainard, D. H. (2004). Color constancy under changes in reflected illumination. *J Vis*, *4*(9), 764–778. http://doi.org/Doi 10.1167/4.9.8

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., … Chichilnisky, E. J. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, *467*(7316), 673–677. http://doi.org/10.1038/nature09424

Finlayson, G. D., Hordley, S. D., & Hubel, P. M. (2001). Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*, 1209–1221. http://doi.org/10.1109/34.969113

Finlayson, G. D., Hubel, P. M., & Hordley, S. (1997). Color by correlation. *Fifth Color Imaging Conference: Color Science, Systems, and Applications*, 6–11. Retrieved from <Go to ISI>://000071602400002

Finlayson, G. D., Mackiewicz, M., Hurlbert, A. C., Pearce, B., & Crichton, S. (2014). On calculating metamer sets for spectrally tunable LED illuminators. *J. Opt. Soc. Am A.*, *31*(7), 10.

Finlayson, G. D., & Morovic, P. (2005). Metamer sets. *J Opt Soc Am A Opt Image Sci Vis*, *22*(5), 810–819. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15898540

Foster, D. H. (2011). Color constancy. *Vision Res*, *51*(7), 674–700. http://doi.org/DOI 10.1016/j.visres.2010.09.006

Foster, D. H., Amano, K., & Nascimento, S. M. C. (2006). Color constancy in natural scenes explained by global image statistics. *Visual Neuroscience*, *23*(3-4), 341–349.

Foster, D. H., & Nascimento, S. M. C. (1994). Relational color constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society of London Series B-Biological Sciences*, *257*(1349), 115–121. Retrieved from <Go to ISI>://A1994PD67000003

Foster, D. H., Nascimento, S. M. C., & Amano, K. (2005). Information limits on identification of natural surfaces by apparent colour. In *Perception* (Vol. 34, pp. 1003–1008). http://doi.org/10.1068/p5181

Foster, D. H., Nascimento, S. M. C., Craven, B. J., Linnell, K. J., Cornelissen, F. W., & Brenner, E. (1997). Four issues concerning colour constancy and relational colour

constancy. *Vision Res*, *37*(10), 1341–1345. Retrieved from <Go to ISI>://A1997WY44900010

García-Pérez, M. a. (1998). Forced-choice staircases with fixed step sizes: Asymptotic and small-sample properties. *Vision Research*, *38*, 1861–1881. http://doi.org/10.1016/S0042-6989(97)00340-4

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual Review of Neuroscience*, *26*, 181–206. http://doi.org/DOI 10.1146/annurev.neuro.26.041002.131116

Gijsenij, A., Gevers, T., & van de Weijer, J. (2011). Computational Color Constancy: Survey and Experiments. *Ieee Transactions on Image Processing*, *20*(9), 2475–2489. http://doi.org/Doi 10.1109/Tip.2011.2118224

Golz, J., & MacLeod, D. I. A. (2002). Influence of scene statistics on colour constancy. *Nature*, *415*, 637–640. http://doi.org/10.1038/415637a

Granzier, J., & Gegenfurtner, K. R. (2011). Memory colour improves colour constancy for unknown coloured objects. *J Vis*, *11*(11).

Granzier, J. J., & Gegenfurtner, K. R. (2012). Effects of memory colour on colour constancy for unknown coloured objects. *Iperception*, *3*(3), 190–215. http://doi.org/10.1068/i0461

Granzier, J., Vergne, R., & Gegenfurtner, K. R. (2014). The effects of surface gloss and roughness on color. *J Vis*, *14*(2), 1–20.

Hansen, T., Giesel, M., & Gegenfurtner, K. R. (2008). Chromatic discrimination of natural objects. *J Vis*, *8*(1), 2 1–19. http://doi.org/10.1167/8.1.2

Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal context on color categories and color constancy. *J Vis*, *7*(4). http://doi.org/Artn 2 Doi 10.1167/7.4.2

Heasly, B. S., Cottaris, N. P., Lichtman, D. P., Xiao, B., & Brainard, D. H. (2014). RenderToolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research. *Journal of Vision*, *14*, 1–22. http://doi.org/10.1167/14.2.6

Hernandez-Andres, J., Romero, J., Nieves, J. L., & Lee Jr., R. L. (2001). Color and spectral analysis of daylight in southern Europe. *J Opt Soc Am A Opt Image Sci Vis*, *18*(6), 1325–1335. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11393625

Heywood, C. a, Gadotti, a, & Cowey, a. (1992). Cortical area V4 and its role in the perception of color. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *12*(October), 4056–4065.

Hunt, R. W. G. (1957). *The reproduction of colour*. London,: Fountain Press.

Hunt, R. W. G. (1991). *Measuring colour*. *Ellis Horwood series in applied science and industrial technology* (2nd ed.). New York: E. Horwood.

Hunt, R. W. G., & Pointer, M. R. (2011). *Measuring Colour: Fourth Edition*. *Measuring Colour: Fourth Edition*. http://doi.org/10.1002/9781119975595

Hurlbert, A. C. (1989). *The computation of color*. Massachusetts Institute of Technology. Retrieved from http://dspace.mit.edu/handle/1721.1/13990

Hurlbert, A. C. (1996). Colour vision: Putting it in context. *Current Biology*, *6*(11), 1381–1384. Retrieved from <Go to ISI>://A1996VT11900013

Hurlbert, A. C. (1997). Primer - Colour vision. *Current Biology*, *7*(7), R400–R402. Retrieved from <Go to ISI>://A1997XK26500006

Hurlbert, A. C. (1998). Computational models of colour constancy. In V. Walsh Kulikowski, J. (Ed.), *Perceptual Constancy: Why things look as they do.* (pp. 283–321). UK: Cambridge University Press.

Hurlbert, A. C. (1999). Colour vision: Is colour constancy real? *Current Biology*, *9*(15), R558–R561. Retrieved from <Go to ISI>://000081850800011

Hurlbert, A. C. (2003). Colour vision: Primary visual cortex shows its influence. *Current Biology*, *13*(7), R270–R272. http://doi.org/Doi 10.1016/S0960-9822(03)00198-2

Hurlbert, A. C., & Ling, Y. Z. (2006). Contextual effects of familiar object colours on collour perception. *Perception*, *35*, 23. Retrieved from <Go to ISI>://000243599300072

Hurlbert, A. C., Pearce, B., Mackiewicz, M., & Finlayson, G. D. (2014). Immersive Colour Constancy by Global Illumination Change Discrimination. *Journal of Vision*, *14*(10), 794.

Hurlbert, A. C., & Wolf, K. (2003). Color contrast: a contributory mechanism to color constancy. *Roots of Visual Awareness*, *144*, 147–160. http://doi.org/Doi 10.1016/S0079-6123(03)14401-0

Hurlbert, A. C., & Wolf, K. (2004). Color contrast: a contributory mechanism to color constancy. *Progress in Brain Research*, *144*, 147–160. http://doi.org/10.1016/S0079-6123(03)14401-0

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, *64, Part 1*(6), 384–404. http://doi.org/10.1037/h0041403

Ives, H. (1912). The Relation Between the Color of the Illuminant and the Color of the Illuminated Object. *Trans. Illuminat. Eng. Soc.*, *7*, 62–72.

Jakob, W., Arbree, A., Moon, J. T., Bala, K., & Marschner, S. (2010). A radiative transfer framework for rendering materials with anisotropic structure. *ACM Transactions on Graphics*. http://doi.org/10.1145/1833351.1778790

Jameson, D. (1983). Some misunderstandings about color perception, color mixture and color measurement. *Leonardo*, *16*, 41–42.

Jobson, D. J., Rahman, Z. U., & Woodell, G. A. (1997). Properties and performance of a center/surround retinex. *Ieee Transactions on Image Processing*, *6*(3), 451–462. Retrieved from <Go to ISI>://A1997WL07500009

Judd, D. B. (1940). Hue, saturation, and lightness of surface colors with chromatic illumination. *Journal of Research of the National Bureau of Standards*. http://doi.org/10.6028/jres.024.016

Kanematsu, E., & Brainard, D. H. (2013). No measured effect of a familiar contextual object on colour constancy. *Color Research and Application*, *00*(00), 1–13.

Kärhä, P., Vaskuri, A., Baumgartner, H., Andor, G., & Ikonen, E. (2013). Relationships between Junction Temperature, Forward Voltage and Spectrum of LEDS. In *CIE Centenary Conference*. Paris.

Kinnear, P. R., & Sahraie, A. (2002). New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5-22 and for age decades 30-70. *Br J Ophthalmol*, *86*(12), 1408–1411. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12446376

Kleiner, M., Brainard, D. H., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception 36 ECVP Abstract Supplement*, 14. http://doi.org/10.1068/v070821

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proc Natl Acad Sci U S A*, *96*(1), 307–312. Retrieved from <Go to ISI>://000078004400056

Krauskopf, J., & Gegenfurtner, K. (1992). Color discrimination and adaptation. *Vision Res*, *32*(11), 2165–2175. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1304093

Krieger, A., Dubin, H., Pearce, B., Aston, A., Hurlbert, A. C., Brainard, D. H., & Radonjic, A. (2014). Illumination discrmination depends on scene surface ensemble. *VSS 2014*.

Land, E. H. (1977). Retinex Theory of Color-Vision. *Scientific American*, *237*(6), 108–&. Retrieved from <Go to ISI>://A1977EC72500006

Land, E. H. (1986). Recent Advances in Retinex Theory. *Vision Res*, *26*(1), 7–&. Retrieved from <Go to ISI>://A1986A509200003

Lee, B. B. (2014). Color coding in the primate visual pathway: a historical view. *J. Opt. Soc. Am A.*, *31*(4), A103–111.

Lin, K. C. (2010). Approach for optimization of the color rendering index of light mixtures. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *27*, 1510–1520. http://doi.org/10.1364/JOSAA.27.001510

Ling, Y. Z. (2005). *The Colour Perception of Natural Objects: Familiarity, Constancy and Memory*. Newcastle University.

Ling, Y. Z., & Hurlbert, A. C. (2008a). Role of color memory in successive color constancy. *J Opt Soc Am A Opt Image Sci Vis*, *25*(6), 1215–1226. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18516130

Ling, Y. Z., & Hurlbert, A. C. (2008b). Role of color memory in successive color constancy. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *25*(6), 1215–1226. Retrieved from <Go to ISI>://000257156000001

Linnell, K. J., & Foster, D. H. (1996). Dependence of relational colour constancy on the extraction of a transient signal. *Perception*, *25*, 221–228. http://doi.org/10.1068/p250221

Linnell, K. J., & Foster, D. H. (2002). Scene articulation: dependence of illuminant estimates on number of surfaces. *Perception*, *31*(2), 151–159. http://doi.org/Doi 10.1068/P03sp

Logvinenko, A. (2015). The geometric structure of color, *15*, 1–9. http://doi.org/10.1167/15.1.16.doi

Logvinenko, A., Funt, B., & Godau, C. (2014). Metamer mismatching. *IEEE Transactions on Image Processing*, *23*(1), 34–43. http://doi.org/10.1109/TIP.2013.2283148

MacAdam, D. L. (1942). visual sensitivities to color differences in daylight. *J Opt Soc Am*, *32*(5), 247–274.

MacDonald, L. W., & Roque, T. (2013). Chromatic Adaptation in an Immersive Viewing Environment. *(Proceedings) 12th Congress of the International Colour Association (AIC), (eds.)*, 623–626.

Mackiewicz, M., Crichton, S., Newsome, S., Gazerro, R., Finlayson, G. D., & Hurlbert, A. C. (2012). Spectrally tunable led illuminator for vision research. *Proceedings of the 6th Colour in Graphics, Imaging and Vision (CGIV), Amsterdam, Netherlands*.

Mackiewicz, M., Pearce, B., Crichton, S., Finlayson, G. D., & Hurlbert, A. C. (2012). Achromatic adjustment outdoors using MEMS reflective display. *Perception*, *41*(12), 1522. Retrieved from <Go to ISI>://000316588600040

Maloney, L. T., & Schirillo, J. A. (2002). Color constancy, lightness constancy, and the articulation hypothesis. *Perception*, *31*(2), 135–139. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11922127

McDermott, K. C., & Webster, M. A. (2012). Uniform color spaces and natural image statistics. *J Opt Soc Am A Opt Image Sci Vis*, *29*(2), A182–7. http://doi.org/10.1364/JOSAA.29.00A182

Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology*, *359*(1), 381–400.

Murray, I. J., Daugirdiene, A., Vaitkevicius, H., Kulikowski, J. J., & Stanikunas, R. (2006). Almost complete colour constancy achieved with full-field adaptation. *Vision Research*, *46*, 3067–3078. http://doi.org/10.1016/j.visres.2006.03.011

Nascimento, S. M. C., Ferreira, F. P., & Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *19*, 1484–1490. http://doi.org/10.1364/JOSAA.19.001484

Nascimento, S. M. C., & Foster, D. H. (2000). Relational color constancy in achromatic and isoluminant images. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *17*(2), 225–231. Retrieved from <Go to ISI>://000085065700004

Novak, C. (1991). Supervised Color Constancy for Machine Vision. *Proceedings of the SPIE: Conference on Visual Processing and Digital Display*.

Ogura, A., Ikeo, K., & Gojobori, T. (2004). Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. *Genome Research*, *14*, 1555–1561. http://doi.org/10.1101/gr.2268104

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *J Vis*, *9*(12). http://doi.org/Artn 6 Doi 10.1167/9.12.6

Parker, A. R. (1998). Colour in Burgess Shale animals and the effect of light on evolution in the Cambrian. *Proceedings of the Royal Society B: Biological Sciences*, *265*, 967–972. http://doi.org/10.1098/rspb.1998.0385

Parraga, C. A., Troscianko, T., & Tolhurst, D. J. (2005). The effects of amplitude-spectrum statistics on foveal and peripheral discrimination of changes in natural images, and a multi-resolution model. *Vision Res*, *45*(25-26), 3145–3168. http://doi.org/DOI 10.1016/j.visres.2005.08.006

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G. D., & Hurlbert, A. C. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PLoS ONE*, *9*. http://doi.org/10.1371/journal.pone.0087989

Pearce, B., Radonjic, A., Dubin, H., Cottaris, N. P., Mackiewicz, M., Finlayson, G. D., … Hurlbert, A. C. (2014). Illumination Discrimination Reveals "Blue" Bias of Colour Constancy in Real and Simulated Scenes. *J Vis*, *14*(10), a599.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*, 437–442. http://doi.org/10.1163/156856897X00366

Rayleigh, L. (1881). Experiments on Colour. *Nature*, *25*(629), 64–66. http://doi.org/10.1038/025064a0

Reeves, A. J., Amano, K., & Foster, D. H. (2008). Color constancy: phenomenal or projective? *Perception & Psychophysics*, *70*, 219–228. http://doi.org/10.3758/PP.70.2.219

Regan, B. C., Julliot, C., Simmen, B., Vienot, F., Charles-Dominique, P., & Mollon, J. D. (2001). Fruits, foliage and the evolution of primate colour vision. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *356*(1407), 229–283. Retrieved from <Go to ISI>://000167971900001

Rushton, W. A. H. (1972). Review Lecture. Pigments and signals in colour vision. *Journal of Physiology*, *220*(3), 1P–31P. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1331666/

Rutherford, M. D., & Brainard, D. H. (2000). The role of illumination perception in color constancy. *Investigative Ophthalmology & Visual Science*, *41*(4), S525–S525. Retrieved from <Go to ISI>://000086246702871

Sandor, N., & Schanda, J. (2006). Visual colour rendering based on colour difference evaluations. *Lighting Research and Technology*. http://doi.org/10.1191/1365782806lrt168oa

Schanda, J., & International Commission on Illumination. (2007). *Colorimetry : understanding the CIE system*. Hoboken, N.J.: Wiley.

Sheats, J. R., Antoniadis, H., Hueschen, M., Leonard, W., Miller, J., Moon, R., … Stocking, A. (1996). Organic Electroluminescent Devices. *Science*. http://doi.org/10.1126/science.273.5277.884

Smith, L., Söderbärg, A., & Björkengren, U. (1994). Continuous ink-jet print head utilizing silicon micromachined nozzles. *Sensors and Actuators A: Physical*. http://doi.org/10.1016/0924-4247(93)00707-B

Smithson, H. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *360*(1458), 1329–1346. http://doi.org/DOI 10.1098/rstb.2005.1633

Smithson, H., & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and levels of reference. *J Vis*, *4*(9), 693–710. http://doi.org/Doi 10.1167/4.9.3

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*, 1711–1737. http://doi.org/10.1016/S0042-6989(00)00021-3

Stockman, A., & Sharpe, L. T. (2006). PHYSIOLOGICALLY-BASED COLOUR MATCHING FUNCTIONS. In *ISS/CIE Expert Symposium '06* (pp. 13–20).

Sumner, P., & Mollon, J. D. (2000). Catarrhine photopigments are optimized for detecting targets against a foliage background. *Journal of Experimental Biology*, *203*(13), 1963–1986. Retrieved from <Go to ISI>://000088329100002

Thomas, P. B. M., & Mollon, J. D. (2004). Modelling the Rayleigh match. *Visual Neuroscience*, *21*(3), 477–482. http://doi.org/10.1017/S095252380421344X

Van De Ven, A. P., Chan, W. K., & Wah, H. C. (2014). LIGHTING DEVICES COMPRISING SOLID STATE LIGHT EMITTERS. US Patent Office.

Vrhel, M. J., & Trussell, H. J. (2002). Color printer characterization in MATLAB. *Proceedings. International Conference on Image Processing*, *1*. http://doi.org/10.1109/ICIP.2002.1038059

Vurro, M., Ling, Y. Z., & Hurlbert, A. C. (2013). Memory color of natural familiar objects: effects of surface texture and 3-D shape. *Journal of Vision*, *13*, 20. http://doi.org/10.1167/13.7.20

Webster, M. A., Mizokami, Y., & Webster, S. M. (2007). Seasonal variations in the color statistics of natural images. *Network*, *18*(3), 213–233. http://doi.org/10.1080/09548980701654405

Werner, A. (2003). The spatial tuning of chromatic adaptation. *Vision Research*, *43*, 1611–1623. http://doi.org/10.1016/S0042-6989(03)00174-3

Werner, A., Sharpe, L. T., & Zrenner, E. (2000). Asymmetries in the time-course of chromatic adaptation and the significance of contrast. *Vision Res*, *40*(9), 1101–1113. http://doi.org/Doi 10.1016/S0042-6989(00)00012-2

Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America a-Optics Image Science and Vision*, *2*(7), 1014–1026. http://doi.org/Doi 10.1364/Josaa.2.001014

Wyszecki, G., & Stiles, W. S. (1982). *Color science : concepts and methods, quantitative data and formulae*. *The Wiley series in pure and applied optics* (2nd ed.). New York: Wiley. Retrieved from http://www.loc.gov/catdir/bios/wiley041/82002794.html

Xiao, B., & Brainard, D. H. (2008). Surface gloss and color perception of 3D objects. *Visual Neuroscience*, *25*(3), 371–385.

Xiao, B., Hurst, B., MacIntyre, L., & Brainard, D. H. (2012). The color constancy of three-dimensional objects. *J Vis*, *12*(4), 6. http://doi.org/10.1167/12.4.6

Xiao, F., DiCarlo, J. M., Catrysse, P. B., & Wandell, B. A. (2002). High dynamic range imaging of natural scenes. In *Tenth Color Imaging Conference: Color Science, Systems, and Applications* (Vol. 2002, pp. 337–342). Retrieved from http://www.ingentaconnect.com/content/ist/cic/2002/00002002/00000001/art00062

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Vision Res*, *41*(20), 2581–2600. Retrieved from <Go to ISI>://000170887600003

Zaidi, Q., & Smithson, H. (2004). Roles for local adaptation and levels of reference in colour constancy. *Perception*, *33*, 24. Retrieved from <Go to ISI>://000224198700076

Zeki, S. (1980). The representation of colours in the cerebral cortex. *Nature*, *284*, 412–418. http://doi.org/10.1038/284412a0

Appendix 1.

The RS5b luminaires came with a stock application to interface with them. This interface did not allow real-time control; furthermore, the basis-functions stored within the program to perform least-squares fitting were hard-coded, and could not be changed after calibration.

Two tools were developed. Firstly, a GUI was developed that replicated the functionality of the stock software, with communications, spectral fitting as outlined in the main text, and colorimetric visualisation to allow fast debugging of spectra that were produced out of calibration. A screenshot of this software is shown below in Figure A1.1. Secondly, the comprising libraries that were developed for the GUI were used with libraries developed for the PR650 spectroradiometer to send values to the lamps (16 bit unsigned integers) and read back the spectral emission, as outlined in Chapter 2; an example procedure for calibration can be seen in Figure A1.2.

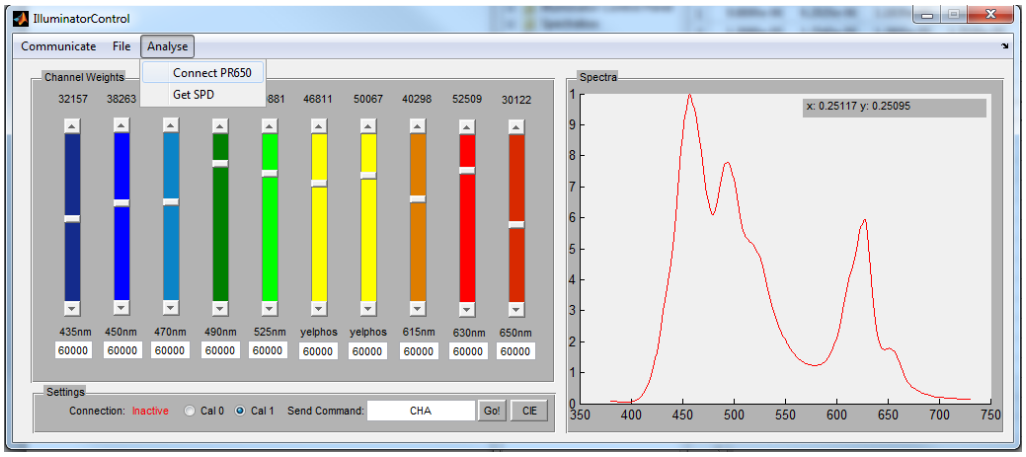

Figure A1.1. Screenshot of the GUI toolbox written in MATLAB to control the luminaires.

```
%% INPUT : MATRIX, a set of 10xN weights
%% OUTPUT: Spectra, 380:4:780nm for each set of weights.

%% Setup Hardware
display('Setting up the Illuminator...')
%spectra_init;

illuminatorConnect(1);
%% wait and warn
display('Reading will start in 30 seconds');
pause(30);
%% Generate a tone.
t1=1/10000:1/10000:0.5;
solt1=(sin(2*pi*196*t1));
sound(solt1,50000);
%% Main loop
display('Reading values')
[l,w] = size(matrix);
percy = (100/w);%(length(matrix)));
for i = 1:w%length(matrix);
    %chw(matrix(i,1:10));
     chw(matrix(:,i)');
    illuminatorRelease;
    display('taking reading...')
    try
        ANS=cal_read;
        Yxy(1) = ANS.bigY;
        Yxy(2) = ANS.x;
        Yxy(3) = ANS.y;
        sp = cal_spectrum;
        pause(15); % To make sure we finish taking measurements
    catch
       spectra_init;
       ANS=cal_read;
       Yxy(1) = ANS.bigY;
        Yxy(2) = ANS.x;
        Yxy(3) = ANS.y;
        sp = cal_spectrum;
        pause(15); % To make sure we finish taking measurements
    end
    display('finished reading...')
    spectrum(:,1) = sp.spectrum(:,1);
    spectrum(:,i+1) = sp.spectrum(:,2);
    save List;
    list(i,:) = [Yxy(1:3)];
    display([int2str(percy*i),'% done']);
    %fclose(Sdev)
    if isunix
        illuminatorConnect(4);
    else
    illuminatorConnect(1);
    end
    save latestMeasurements;

    %chw([0 0 0 0 0 0 0 0 0 0])
    %pause(1)

end
save List;
save workspace;
illuminatorConnect(0);
```

Figure A1.2. Simple code procedure for calibrating the illuminator; the weights sent could be maximal channel values or arbitrary units.

Appendix 2.

Table A2.1. CIE 1931 Yxy coordinates of the fruit, and chromaticity matched papers produced with the calibrated Inkjet printer; measurements taken under D67.

	Fruits		Papers		
	x	**y**	**x**	**y**	**ΔE_{uv}**
Pear Reading 1	0.302	0.341	0.301	0.362	8.58
Pear Reading 2	0.352	0.43	0.315	0.406	13.98
Pear Reading 3	0.318	0.365	0.317	0.399	12.61
Pear Reading 4	0.304	0.34	0.301	0.362	9.42
Banana Reading 1	0.447	0.458	0.447	0.458	0.00
Banana Reading 2	0.434	0.457	0.434	0.457	0.00
Banana Reading 3	0.443	0.465	0.443	0.465	0.00
Apple Reading 1	0.503	0.348	0.499	0.353	4.28
Apple Reading 2	0.506	0.348	0.476	0.353	17.18
Apple Reading 3	0.439	0.352	0.465	0.355	12.13
Mean					7.82

Appendix 3.

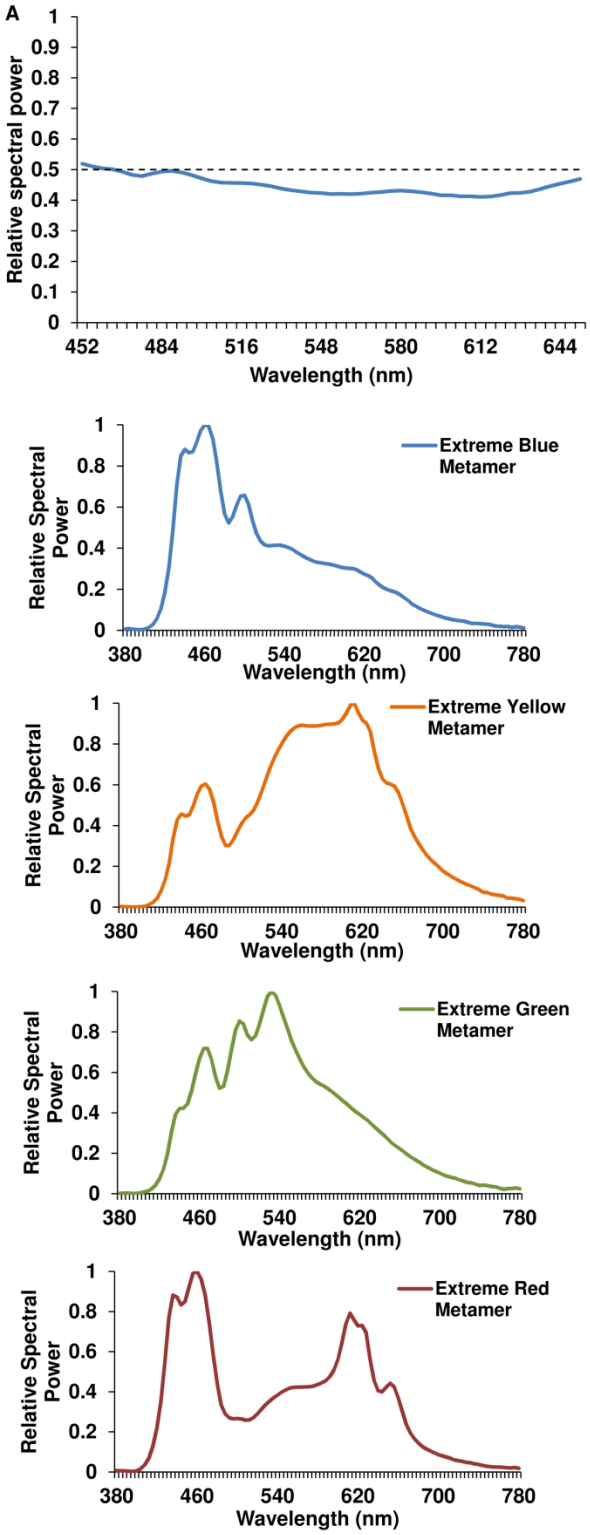

Figure A3.1. Top: surface reflectance of the grey card lining of the viewing box. Below: measured spectra of each of the extreme comparison illuminations. From Pearce et al (2014).

Appendix 4.

Standardised Instructions

The experimenter will ask you to take a seat and comfortably position yourself so that you can see into the viewing box; please get as close as you can.

You will be given a joy pad controller; the experimenter will indicate two buttons, [1] and [2], which will be used in the experiment to provide answers. When the experiment begins you will be shown a light that illuminates the viewing box. Then there will be two subsequent lights, you are asked to signal which is most like the target, using either of the buttons, [1] denoting the first light is most similar, or [2] for the second light.

Each trial will be preceded by two beeps, and then the target light will be presented. A further beep will occur before each comparison light. There are four blocks in the experiment, each block taking around 15 minutes. You will hear 4 beeps at the end of each block; please alert the experimenter when the block has ended. When the block has ended, the experimenter will set up the viewing box with the next condition or the experimenter will schedule a time to continue with subsequent blocks if more convenient with you.

It is the purpose of this experiment to assess your thresholds of discrimination of lights under different conditions.

Please signal the experimenter if you feel that you would like a break, or would like to withdraw from the study; which can be done at any time. Your data will be kept anonymous and will only be used for data analysis purposes as part of a body of data.

If you have any questions please ask the experimenter, either before or at any time after the study; contact details are stated below.

Experimenter:	Bradley Pearce,	b.m.pearce@ncl.ac.uk
Principal Investigator:	Anya Hurlbert,	anya.hurlbert@ncl.ac.uk