



A University of Sussex PhD thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

THE COMPONENTS OF COLOUR VISION

Marie Rosanna Rogers

Thesis submitted for the degree of Doctor of Philosophy

University of Sussex

May 2018

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:

.....

Article Format Thesis Declaration

The thesis conforms to an ‘article format’ in which the first chapter presents an overview of the relevant literature, an outline of the empirical work of the thesis, and discussion of the overall contribution of the thesis to the field. The remaining four chapters consist of discrete papers written for publication in peer-reviewed journals. Two of these chapters have been published, and two are prepared for submission.

Chapters and author contributions

Chapter 1 provides an overview of the field of colour vision and colour perception development, the novel contribution to the field this thesis provides, and a discussion of the findings of the chapters in relation to the existing literature.

Chapter 2 has been written in a style appropriate for *Journal of Experimental Child Psychology*:

Rogers, M. R., Witzel, C. W., Rhodes, P., & Franklin, A. Colour Constancy and Colour Term Acquisition are Positively Related in Early Childhood.

Author contributions: MRR, CW and AF designed the research; PR printed the custom stimuli, MRR collected the data; MRR and AF analysed the data; and MRR, CW, and AF wrote the paper.

Chapter 3 is published in the *Journal of the Optical Society of America A* as:

Rogers, M. R., Knoblauch, K., Franklin, A. (2016). Maximum Likelihood Conjoint Measurement of Lightness and Chroma. *Journal of the Optical Society of America A*. 30(3), A184-A193. doi: 10.1364/JOSAA.33.00A184.

Author contributions: MRR, KK and AF designed the research; MRR collected the data; MRR and KK analysed the data; and MRR, KK, and AF wrote the paper.

Chapter 4 3 is published in the journal *Infancy* as:

Rogers, M. R., Knoblauch, K., Franklin, A. A Novel Method to Investigate the Development of Dimensions during Infancy. *Infancy*. 23(6), 833-856. doi: 10.1111/inf.12260

Author contributions: MRR, KK and AF designed the research; MRR collected the data; MRR and KK analysed the data; and MRR, KK, and AF wrote the paper.

Chapter 5 is written in a style suitable for *Vision Research*:

Rogers, M. R., Lee, R., & Bosten, J. Why is Colour Discrimination Poorest Along the Daylight Locus?

Author contributions: MRR and JB designed the research; RL helped with stimulus production; MRR collected the data; MRR and JB analysed the data; and MRR and JB wrote the paper.

Acknowledgments

The years culminating in this thesis have been wonderful, challenging, complex, rich and varying. There are so many people who have helped me, directly or indirectly, along the path towards the final product. There are not enough opportunities to acknowledge them, so I do not intend to miss this one.

I cannot imagine better supervisors than Professor Anna Franklin and Dr Jenny Bosten. They have been inspirational, stimulating and supportive. From the day of my interview, Anna has helped me with endless patience and vision. Jenny's scientific curiosity and determination made conducting research with her immersive. Conversations with Jenny, on any topic, are mind-stretching and exciting. I am so grateful to you both. I also thank my collaborators: Kenneth Knoblauch, Christoph Witzel and Rob Lee who were generous with their time and taught me a wide array of skills.

The Sussex Colour Group is a special place to work. John Maule and Lewis Forder were always on hand for technical advice and support. Gemma Catchpole could be relied upon to make testing babies and children easier, and assisted with collecting data for Paper 1. Lucy Somers brought her enthusiasm and artistic eye to the lab, plus a willingness to swim in the sea with me after work. Jessica Banks helped with spectacular cakes. Leticia Álvaro provided her sense of humour and coding expertise. Alice Skelton made me laugh every day we shared an office, often to the annoyance of our neighbours.

The friendships I made in Brighton over the course of my PhD gave me an invaluable network of support, and so many joyful memories: Anne, Kate, Jonny, Charlotte, Sam, Tom, Amy and Bea. Thank you. I also thank my Bristol friends: Mathew, Claire, Richard and Christopher. You have provided support, homes to escape to, and new perspectives. And thank you to my dear old friend Amy Brassington for being a constant in my life, and a role model for brilliance and resilience.

Finally, and most importantly, I thank my mum Angela. I cannot convey here how grateful I am for the love and support you have given me. Thank you for encouraging me to be curious.

UNIVERSITY OF SUSSEX

MARIE ROSANNA ROGERS – PHD PSYCHOLOGY

THE COMPONENTS OF COLOUR VISION

SUMMARY

Colour perception is formed of many different components, such as colour discrimination, colour constancy, colour term naming, and the dimensions of colour (hue, chroma and lightness). It is a ‘toolbox’ of processes, not one cohesive function. Some of the components of colour vision develop into adult-like function over childhood, but they do not necessarily mature at the same speed. The studies in this thesis investigate adult, child and infant colour perception and cognition.

Paper 1 finds a relationship between colour constancy and colour term naming in three- to four-year-old children. This relationship has wider implications for the co-development of language and perception. Paper 2 (Rogers, Knoblauch & Franklin, 2016) uses the technique of Maximum Likelihood Conjoint Measurement (MLCM) in adult participants to investigate the interaction between lightness and chroma in perception.

Paper 3 combines MLCM analysis with preferential looking methods to compare interaction of lightness and chroma in infant and adult participants. This study paves the way for the use of MLCM and eye-tracking for studying other dimensions in development such as face perception, language, surface and shape. Paper 4 investigates why discrimination is poorest along the blue-yellow direction of cone opponent space (also known as the daylight locus). We tested the theory that this is adaptive for colour constancy by comparing illumination discrimination to surface discrimination in adult participants. We found equally poor discrimination for blue-yellow in both conditions, suggesting colour constancy is not the only explanatory factor.

Together, these papers add to our understanding of the key components of colour vision over the life span and how perception of colour depends on various contextual and individual factors. Furthermore, this thesis develops novel applications of experimental techniques, and paves the way for these methods to be used to study other cognitive and developmental domains.

Table of contents

Table of contents	vii
Chapter 1: Introduction and Thesis Overview	1
1.1 Colour: From retina to cortex.....	3
Trichromacy and opponent processes theory	3
The evolution of human colour vision	6
Colour in the cortex.....	7
The development of colour mechanisms.....	8
1.2 The dimensions of colour appearance.....	10
Colour appearance models	10
The measurement of colour appearance.....	13
The development of colour appearance	13
1.3 Colour term acquisition.....	15
1.4 Perception of surface and illumination chromaticity	17
Colour constancy.....	22
The development of colour constancy	24
1.5 Thesis overview and research questions	26
Paper 1: Colour Constancy and Colour Term Acquisition are Positively Related in Early Childhood	27
Paper 2: Maximum Likelihood Conjoint Measurement of Lightness and Chroma	29
Paper 3: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy	31
Paper 4: Why is Colour Discrimination Poorest along the Daylight Locus?.....	33
1.6 Overall contribution	36
Contribution to understanding of the development of colour perception	36
Contribution to understanding the dimensions of colour appearance.....	36
Contribution to understanding colour constancy and colour discrimination	37

1.7 Future research	39
1.8 Conclusion	43
Chapter 2: Paper 1: Colour Constancy and Colour Term Knowledge are Positively Related in Early Childhood	45
2.1 Abstract	46
2.2 Introduction	47
2.3 Methods	52
Participants	52
Colour constancy task	53
Colour naming task	64
2.4 Results	65
2.5 Discussion	71
Chapter 3: Paper 2: Maximum Likelihood Conjoint Measurement of Lightness and Chroma	75
3.1 Abstract	76
3.2 Introduction	77
The dimensions of colour appearance	77
Conjoint measurement	79
The current study	80
3.3 Experiment 1: Methods	83
Observers	83
Apparatus	83
Stimuli	83
Design and Procedure	87
3.4 Experiment 1: Model	88
3.5 Experiment 1: Results	89
Interpolation	91
3.6 Experiment 2: Methods	93

Observers	93
Apparatus	93
Stimuli	93
Design and Procedure	95
3.7 Experiment 2: Model	96
3.8 Experiment 2: Results	98
3.9 Discussion	106
Chapter 4: Paper 3: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy	112
4.1 Abstract	113
4.2 Introduction	114
Perceptual salience of physical dimensions	114
Interaction of dimensions	115
Maximum Likelihood Conjoint Measurement	117
Current study	118
4.3 Method	121
Participants	121
Stimuli and apparatus	121
Design and procedure	124
4.4 Data analysis	129
4.5 Results	133
Infants	133
Adults	140
4.6 Discussion	144
Development of dimensional interaction	145
Future directions	148
Conclusion	149

Chapter 5: Paper 4: Why is colour discrimination poorest along the daylight locus? ..	150
5.1 Abstract	151
5.2 Introduction	152
Colour constancy and the daylight locus	153
The range-accuracy trade-off hypothesis	157
Current study	158
5.3 Methods.....	161
Participants	161
Design	161
Apparatus	161
Stimuli	162
Procedure	168
5.4 Data Analysis	170
5.5 Results	172
5.6 Discussion	178
Conclusion	180
References	183

Chapter 1: Introduction and Thesis Overview

Our experience of colour is holistic. That is, perception of colour in our day-to-day life is functional and it is not useful to think about its constituent parts. However, colour perception is comprised of a toolbox of components coded by the brain in various ways. For example, the separate neural encoding of colour dimensions (hue, saturation and chroma; Shapley & Hawken, 2011), colour categories (Bird, Berens, Horner, & Franklin, 2014), and the perception of illumination and surface chromaticity (Roe et al., 2012). Learning to integrate, or separate, the perceptual components of colour also develops over childhood at different rates. For example, there is evidence that colour constancy matures during early childhood (Dannemiller & Hanko, 1987; Witzel, Sanchez-Walker, & Franklin, 2013) and the dimensions of colour appearance integrate over development (Kaldy & Blaser, 2009). Context and experience may affect our perception of colour, such as the background to the object (Chubb, Sperling, & Solomon, 1989), the illumination (Arend & Reeves, 1986), our cultural expectations (Hurlbert & Ling, 2005), memory demands (Allred & Olkkonen, 2015; Olkkonen & Allred, 2014) and our prior experience (Mollon, Bosten, Peterzell, & Webster, 2017; Witzel, Racey, & O'Regan, 2017).

Colour pervades our experience of the world, but it can be difficult to define and describe. Although colour *appearance* is subjective, colour itself has a measurable, physical basis in the illumination and reflection of light. This allows us to conduct experiments that compare the physical basis of colour to its perception. This thesis investigates some of the components of colour perception in infancy, childhood and adulthood. As colour perception adapts to the current task or demand, it is important to consider all components of colour perception. Whereas prior work tended to focus on

individual components of colour perception in isolation, the current thesis considers interactions and maturation of the components of colour vision.

The overall aim of this thesis is to investigate some of the key components of colour vision in child development and adulthood. This chapter introduces the components of colour vision that are investigated in the further chapters. There will be a brief introduction to the physiological and neurological coding of colour in the visual system, followed by a discussion of the dimensions of colour appearance and the perception of surface change and illumination change. There will then be an overview of the rationale for the four experimental papers that follow in the thesis, the key research questions of the papers and how the methods of the experiments are designed to answer these questions. Finally, this chapter will detail how the findings presented in this thesis provide a novel contribution to the field, and future directions for the work will be suggested.

1.1 Colour: From retina to cortex

Trichromacy and opponent processes theory

There are three types of photoreceptor in the normal trichromatic human retina. They are sensitive to different wavelengths of light in the visible spectrum. The three types of cell have a peak sensitivity at 564 – 580 nm, 534 – 545 nm and 420 – 440 nm and are designated as long (L), medium (M) and short (S) respectively (V. C. Smith & Pokorny, 1975). Figure 1.1 shows the spectral sensitivities of the three cone types. The Young–Helmholtz trichromatic theory, developed in the 19th Century, states that all colour sensation simply arises from a mixture of three signals (Helmholtz, 1896; Young, 1800, 1802). However, in 1878, Hering developed the *opponent processes theory* of colour perception (Hering, 1874; Turner, 1993). Hering observed that there are perceptually pure colours (red, green, yellow and blue), that produce single colour sensations, which cannot be described as a mixture of any other colours. These “unique hues” are opponent in the sense that it is not possible to imagine, for example a reddish-green or a yellowish-blue. This suggests that the unique hues are at the opposite ends of two intersecting axes: one with red and green at opposite ends, and the other with blue and yellow. At the centre of the intersecting axes is ‘unique white’: a colour neither red, green, blue nor yellow. Figure 1.2 shows the unique hues represented in an opponent format.

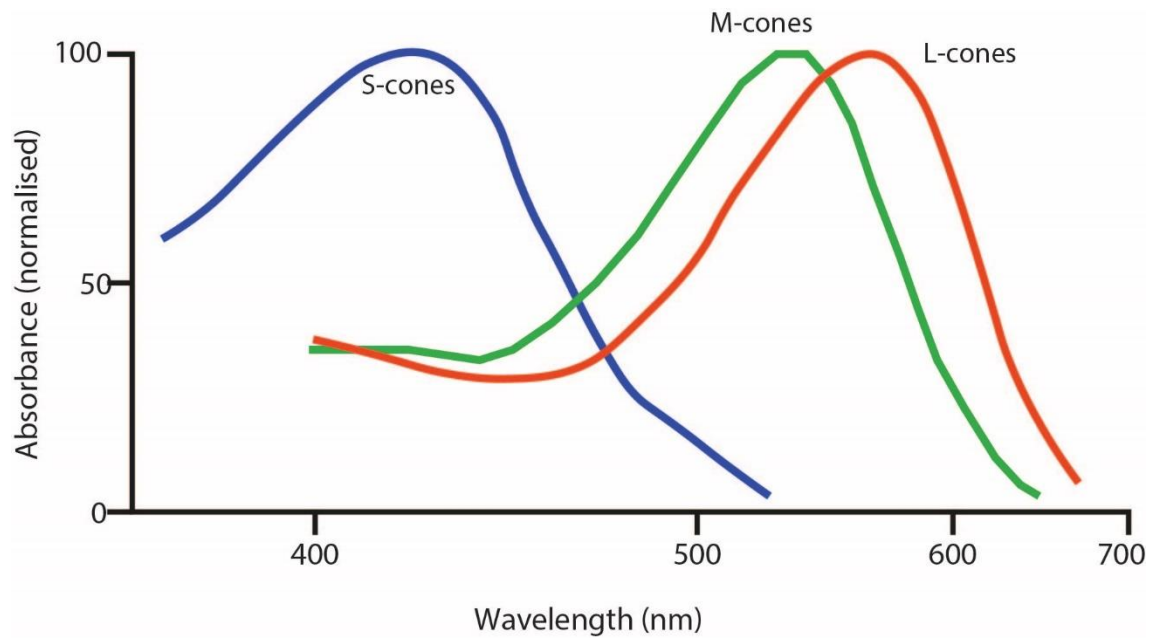


Figure 1.1. The spectral sensitivities of the long- (L-), middle- (M-), and short- (S-) wavelength cones in the retina.

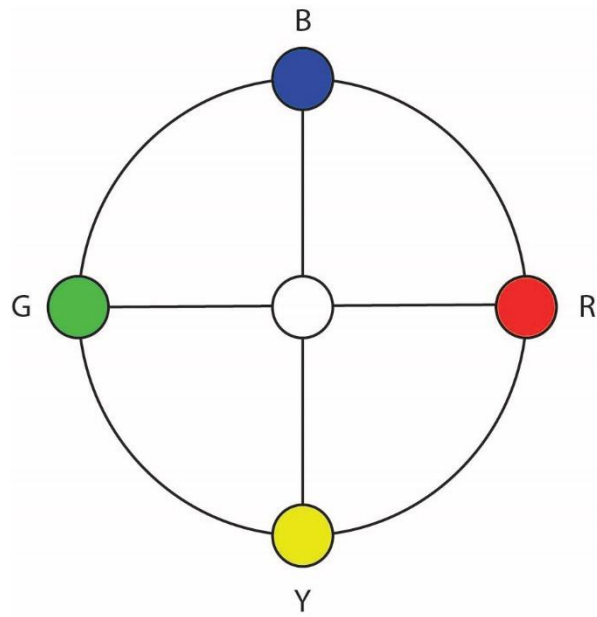


Figure 1.2. Colour opponency: the four unique hues (plus white in the centre) are at the opposite ends of two axes.

Later, in the mid-20th Century, psychophysical studies found support for opponent mechanisms, thus adding weight to the opponent processes theory (Hurvich & Jameson, 1957; Jameson & Hurvich, 1955). It is now thought that the photoreceptor cells in the retina pass signals on to the next stage of visual processing, which compares the activity of the three different cone types in order to compute colour (Mollon & Cavonius, 1987). The opponent mechanisms are computationally important, as there is a large amount of overlap between the cones in the spectral sensitivities (see Figure 1.1), therefore the human visual system must disentangle the high correlation of signals (Buchsbaum & Gottschalk, 1983). There are three separate channels in the retinal ganglion cells that convey information about colour to the brain. Two are chromatic opponent mechanisms, encoding $L / (L + M)$ and $S / (L + M)$. These are also known as the cardinal colour directions. The third mechanism represents luminance intensity and is formed by $L + M$ signals (Derrington, Krauskopf, & Lennie, 1984; Krauskopf, Williams, & Heeley, 1982).

The evolution of human colour vision

The cardinal mechanisms are an example of separate components of colour vision: they evolved at different times in our ancestry, and they are anatomically distinct. Molecular genetic sequencing suggests that short-wavelength sensitive cones became distinct from medium-wavelength sensitive cones (or possibly from ancestral rhodopsin pigment cells), around 500 million years ago, thus forming the basis for dichromatic vision (Nathans, Thomas, & Hogness, 1986). L and M cones became distinct from a general middle-wavelength sensitive (or “yellow”-sensitive; Conway, 2009) cone much later, around 30 million years ago (Nathans et al., 1986). Therefore, in evolutionary terms, the $L / (L + M)$ subsystem, is much more recent than the S-based subsystem (Mollon,

1989). The spectral sensitivity curves of L and M cones are only different by a shift of around 30nm, and they are genetically highly homologous (Nathans et al., 1986).

Humans are amongst only a few species of mammal (alongside a few other primates; Surridge, Osorio, & Mundy, 2003) with trichromatic vision, which suggests there must be a particular, unique environmental pressure to evolve trichromacy (Dominy & Lucas, 2001; Mollon, 1999). There are various theories about why early humans developed trichromacy, including the idea that the L/(L+M) subsystem was useful in seeing ripe fruit amongst foliage (Osorio & Vorobyev, 1996; Regan et al., 2001), and that there is an advantage to recognising skin tones as a cue to health (Benitez-Quiroz, Srinivasan, & Martinez, 2018; Changizi, Zhang, & Shimojo, 2006). The cardinal mechanisms evolved to extract useful information from the natural environment, and therefore are likely adapted to the scene statistics of natural environment. Furthermore, the spectral sensitivities of the cone fundamentals may themselves be attuned to statistics of the natural environment (Osorio & Bossomaier, 1992).

Colour in the cortex

Whilst the retinal structures supporting colour vision are relatively well mapped out, the higher level, cortical basis of colour is still not fully understood (Conway et al., 2010). Colour is an integral aspect of vision, and connections carrying colour information are distributed throughout the brain (Gegenfurtner, 2003). The components of colour appear to be processed more separately at early stages of the visual system, and become integrated at higher levels. For example, there are separate pathways in the lateral geniculate nucleus (LGN; Webster, & Mollon, 1991) where neurons respond linearly to cone input (Derrington et al., 1984), whereas colour and visual form appear to be integrated in area V2 (Gegenfurtner, Kiper, & Fenstemaker, 1996; Shipp & Zeki, 2002).

Although colour processing is distributed, there is also evidence that different brain areas subserve separate encoding of the components of colour perception. In V4, the illumination chromaticity appears to be discounted from the surface reflectance, supporting a stable representation of surface colour across illumination changes, i.e. colour constancy (Heywood, Gadotti, & Cowey, 1992; Kusunoki, Moutoussis, & Zeki, 2006; Roe et al., 2012; Schein & Desimone, 1990; Shipp & Zeki, 1985; Walsh, Carden, Butler, & Kulikowski, 1993). This suggests that the brain forms separate representations of the components surface colour and illumination at this stage. Furthermore, single cell recording in macaques reveals that there are spatially organised retinotopic “hue maps” in V1 in which nearby regions represent colours of a similar hue (Xiao, Wang, & Felleman, 2003). An fMRI study in humans has shown that colour categories, as opposed to metric hue differences, appear to be coded in different brain regions (Bird et al., 2014).

The development of colour mechanisms

At birth, the human retina is still anatomically immature (Abramov et al., 1982; Yuodelis & Hendrickson, 1986). However, even in the first few weeks of life there are rapid developments in the visual system. For example, the photopigments in the cone cell elongate, which impacts on the spectral sensitivities of the receptors. Furthermore, photosensitive cells migrate from the fovea to the edges of the retina, whilst the density of colour sensitive pigments in the fovea increases (Abramov et al., 1982). These physiological changes support the rapid development of colour vision in the first few months postnatally.

Experimental studies with infant participants, using techniques such as preferential looking, have demonstrated that two-month-old infants have trichromatic vision

(Bornstein, 1976; Suttle, Banks, & Graf, 2002). Infants do not appear to discriminate between achromatic and blue chromatic stimuli at birth (Brown & Teller, 1989; Zemach & Teller, 2007), suggesting they do not have fully functioning trichromacy. Bornstein (1976) showed that two-month infants could discriminate colour differences that adult dichromats could not. Furthermore, Brown & Teller (1989) tested two-month-old infants' spectral sensitivity and found a dip in discrimination which is characteristic of trichromatic vision. By three to four months of age, infants have adult-like trichromatic vision (Banks & Bennett, 1988; Morrone, Burr, & Fiorentini, 1993; Volbrecht & Werner, 1987), in that the opponent pathways are fully functioning. However, infants do not yet have fully mature chromatic discrimination by this age.

At four months, the ratio of input across the cone types is similar to that of adults, but there is poorer discrimination compared to adults (Knoblauch, Vital-Durand, & Barbur, 2001). The physiological mechanisms supporting colour vision are fully developed, but infants do not necessarily yet have adult-like colour perception. The ability to discriminate colours progressively improves through development until adolescence (Knoblauch et al., 2001). The development of colour perception appears to depend on experience (Sugita, 2004).

1.2 The dimensions of colour appearance

Colour appearance models

So far, this chapter has described the physiological and neural coding of colour in the human visual system, but this does not tell us about the subjective experience of colour (Smithson, 2015). Colour appearance models, or colour spaces, aim to represent the perceptual organisation of colour, as opposed to its physical or physiological organisation. Models often describe the appearance of colour as varying along three dimensions: hue, lightness and saturation (Wyszecki & Stiles, 2000). The CIE 1931 colour spaces were developed by the Commission internationale de l'éclairage (International Commission on Illumination) based on a series of colour matching experiments, in which observers adjusted the values of three primary lights until they produced the same appearance as a reference light (Guild, 1932; Wright, 1929). The results of these experiments allowed the formation of CIE 1931 RGB coordinates, where R, G and B specify the three primary light intensities to match a colour. The CIE 1931 XYZ colour matching functions are a linear transformation of the RGB coordinates. CIE xyY was derived from XYZ as a way to graphically present the chromaticity of colours, where x and y describe the chromaticity and Y describes the luminance.

Perceptual colour spaces aim to describe and predict the appearances of lights and surfaces, so that the metric distances in the colour space are approximately perceptually uniform. For example, the CIELUV and CIELAB colour difference formula were derived in 1976 by the CIE, with the aim of greater perceptual uniformity than xyY (Robertson, 1990). CIELUV is a linear transformation of xyY. CIELAB defines a colour space with one channel for luminance (L) and two colour channels describing

variation in red-green (A) and blue-yellow (B). When converting between perceptual colour spaces, or applying them to different settings (e.g. different monitor screens), it is important to specify the whitepoint. The whitepoint is a set of coordinates corresponding to white, and allows for such calibration.

There are also colour spaces that are based on the physiological organisation of colour perception (as opposed to the perceptual organisation). Derrington-Krauskopf-Lennie (DKL) space is a cone opponent space that uses the three cardinal mechanisms as the axes (Derrington et al., 1984). Derrington et al presented a technique for analysis of the chromatic properties of neurons in the LGN of the macaque. Colour is represented in three dimensions, first along an axis where only luminance varies, second along an axis where chromaticity varies without excitation of short-wavelength sensitive (S) cones, and third along an axis where chromaticity varies without excitation of long- (L) or medium-wavelength sensitive (M) cones (see Figure 1.3). The orthogonal axes intersect at a white point, and the plane of this intersection is called the azimuth. The MacLeod-Boynton (1979) chromaticity diagram is a two-dimensional plane roughly equivalent to the first two (chromatic) dimensions of DKL space and based on the Smith and Pokorny cone fundamentals (V. C. Smith & Pokorny, 1975). It only represents hue and saturation: $(L / L + M)$ varies along one dimension and $(S / L + M)$ varies along the other.

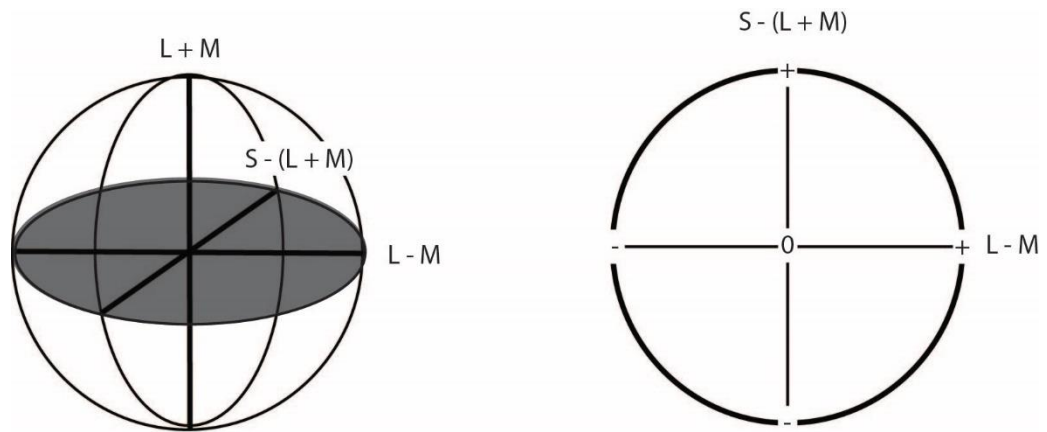


Figure 1.3. Derrington-Krauskopf-Lennie (DKL) space. A colour space based on the chromatic cardinal mechanisms.

The measurement of colour appearance

As outlined above, colour appearance is generally described in the three dimensions of lightness, hue, and saturation. In attempts to produce perceptually uniform colour appearance models, there has been much discussion about the independence of the dimensions of colour appearance (Krantz, 1975; Pieters, 1979). A perceptual dimension is ‘independent’ when changes in that dimension do not change the appearance of a different dimension. However, this is not always the case for the dimensions of colour appearance (Burns & Shepp, 1988). This lack of independence presents challenges for designing colour appearance models. There are various examples where the lack of independence between the dimensions of colour appearance is particularly notable. For example, the Bezold-Brücke effect describes when a change in luminance alone can cause a perceived shift in hue (Walraven, 1961). Similarly, the Abney effect describes when a white light is added to a monochromatic light but there is a change in apparent hue (Abney, 1909), and the Helmholtz-Kohlrausch effect (Nayatani, 1998) describes when chroma contributes to perceived brightness.

Various methods have been used to investigate the interaction of the dimensions of colour appearance. Multidimensional scaling uses similarity or difference judgments or ratings between stimuli to determine the number of dimensions necessary to describe the perceptual space of the stimuli and the distribution of stimuli in that space (e.g. Bimler, 2011; Burns & Shepp, 1988; Ekman, 1954; Helm, 1964; Indow, 1988; Indow & Kanazawa, 1960).

The development of colour appearance

The development of the low-level physiological structures supporting colour vision are relatively well understood, as outlined in section 1.1. However, there are many

unanswered questions about the subjective experience of colour during development. For example, is one dimension of colour appearance particularly salient for infants (A. M. Brown & Lindsey, 2013)? Are we born with the ability to see colour as a unified whole, or do we integrate colour dimensions over the course of development (Káldy & Kovács, 2003)? Do infants organise colour into discrete categories, as adults do (Skelton, Catchpole, Abbott, & Franklin, 2017)?

The integration of perceptual dimensions is an important theme of developmental research (L. B. Smith, 1980). There is evidence that the dimensions of colour appearance are processed differently in infants compared to adults; for example, six-month-old infants appear to ignore lightness differences in preferential looking tasks (A. M. Brown & Lindsey, 2013). It may be the case that adult-like processing of colour appearance is developed by learning about the way attributes covary in natural environments. For example colour and luminance often covary due to changes in illumination in a scene (Golz & MacLeod, 2002), which may help infants to individuate objects (Woods & Wilcox, 2010).

As adults, we tend to categorise colours into discrete groups, (such as red, blue, green, yellow, etc.) depending on our native language (Roberson, 2005). There is debate about the extent to which these categories are universal and innate, versus culturally-dependent and learned (Bird et al., 2014; Bornstein, Kessen, & Weiskopf, 1976; Clifford et al., 2012; Roberson, Davies, & Davidoff, 2000). Recent evidence suggests that infant categories are aligned with the mechanisms of colour vision, and that the cone-opponent mechanisms provide the fault lines around which the adult colour lexicon is formed (Skelton et al., 2017).

1.3 Colour term acquisition

Many researchers have noted that children find it more difficult to learn colour terms than other types of words (e.g. Bornstein, 1985; Soja, 1994). At three years of age, for example, children may know thousands of nouns for very specific things, but they still cannot apply colour words in an adult manner. This is in contrast to the observation that children are quick at picking up the meanings of object words, sometimes after just one exposure (Heibeck & Markman, 1987). Rate of colour term acquisition varies greatly across children. Kowalski and Zimilies (2006) found in a sample of 67 children aged between 25 – 39 months that a third could name no colour terms whereas a third could name 5 or more. It is unclear where this variation comes from: what makes some children quicker colour term learners than others?

Various theories have been proposed to account for the delay in colour term acquisition. For example, one theory suggests that colour is an abstract property of objects, and thus a difficult concept for children to navigate (Soja, 1994). Some have argued that the word order of colour naming in English constrains learning, and that children's colour term learning in languages without this property is not so delayed (Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). Researchers have also noted that colour is a categorical construct and thus children must learn the category boundaries agreed upon by their native language before applying the words consistently (Wagner, Dobkins, & Barner, 2013). Furthermore, learners of some languages may have to modify pre-existing perceptual categories present in infancy if they don't align with those in their own colour lexicon (Bornstein et al., 1976; Franklin & Davies, 2004; Skelton et al., 2017). One possibility is that colour term learning is constrained by perceptual limitations, specifically, colour constancy development. Colour term knowledge may help 'anchor' the representation of the surface colour during illuminant changes via either implicit or

explicit top-down influence of colour naming (Witzel, Maule, & Franklin, 2013; Witzel, Sanchez-Walker, et al., 2013). It is likely that multiple constraints act upon children's ability to learn colour terms (Franklin, 2006) and that level of colour constancy development cannot entirely explain the delay.

1.4 Perception of surface and illumination chromaticity

Light reaches the retina directly from an illumination source, or as a reflection from a surface. Light enters the eye as electromagnetic energy and thus possesses a spectral power distribution, i.e. a wavelength-by-wavelength distribution of energy in the visible region of wavelengths. Figure 1.4 shows A) a specular reflection and B) a diffuse reflection. A smooth or glossy surface directly reflects light to a greater extent than a matte surface. A mirrored surface reflects the illumination equal and opposite to the angle of incidence with no subsurface interactions that alter the reflected spectrum.

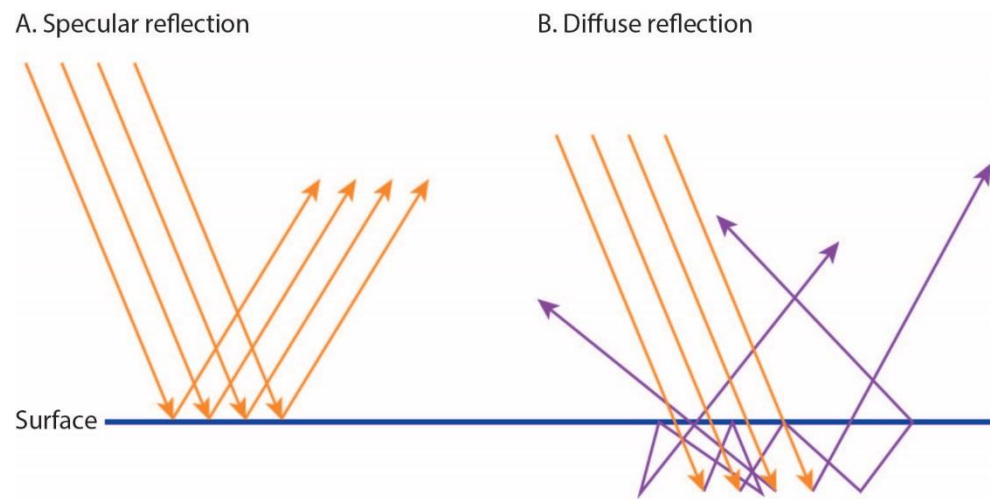


Figure 1.4. The components of reflected light. A) shows specularly reflected light and B) shows diffuse reflection.

A *specular highlight* is the point on a glossy/smooth surface where the illumination spectrum is predominantly reflected. This may be used by the visual system as a cue to the illumination (see Colour constancy section below). Diffuse light penetrates the object's surface, interacts with subsurface pigment molecules (thus altering its spectrum) and is then re-emitted diffusely at a range of angles. Specularly reflected light from an object has the same spectrum as the illumination. Diffusely reflected light from an object has a spectrum that is the product of the illumination spectrum and the surface reflectance spectrum (Smithson, 2005). See Equation 1 below, where λ is a function of wavelength of visible light, from 400 to 700 nm (i.e. a spectrum), E is the illumination and R is the surface reflectance.

$$E(\lambda)\{R_{body}(\lambda) + R_{specular}(\lambda)\} = \text{reflected light}$$

Equation 1.1.

It is unclear to what extent we can maintain separate concurrent representations of surface and illumination colour. Illumination and surface chromaticity representation is an example of the separate components of colour vision working together to solve a perceptual task. The perception of transparency illustrates how illumination and surface perception work in tandem. Transparency occurs through different localised illuminations through and around a filter; the illumination's spectrum is changed by the filter. Figure 1.5 shows a 'transparency illusion' in which two coloured squares appear to overlap. The image creates the illusion of a 'greenish-red' or a 'reddish-green' in the overlapping section, despite red and green being unique hues. Unique hues have been defined as single, pure perceptual experience; it is impossible to envision a "bluish-yellow" or a "reddish-green" appearance. This observation helped form the theory of opponent processes (Helmholtz, 1896; Young, 1800), which underlies our

understanding of colour vision. According to a simple opponent model of colour appearance, it should not be possible to perceive blue and yellow at once at the same spatial location. However, an illusion of transparency *can* lead to perception of two opponent colours simultaneously: colour vision can represent independent colours for the same patch of retina as separate components (Ekroll, 2005).

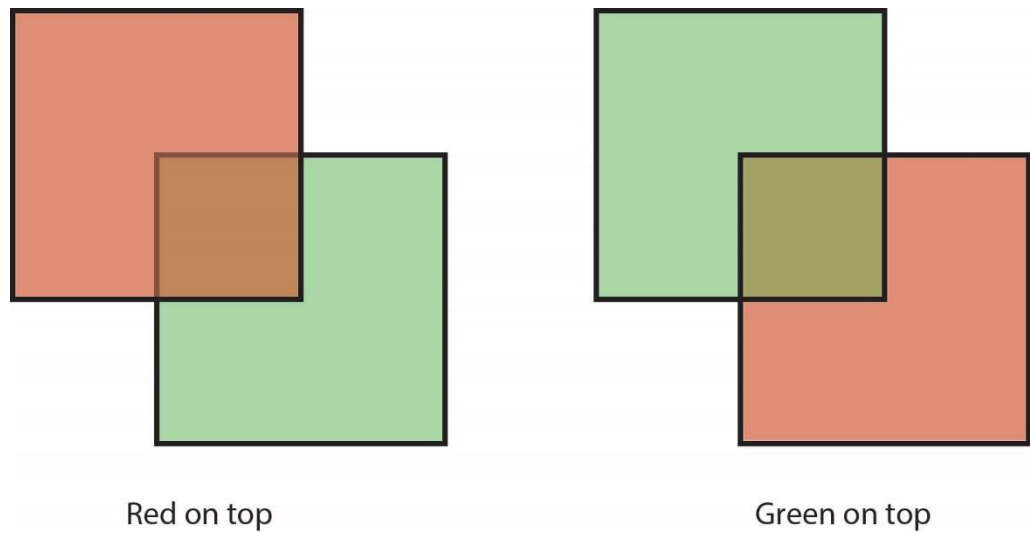


Figure 1.5. Transparency illusion. Although the definition of unique hues means that we cannot conceive of a “reddish-green” or a “greenish-red”, the illusion of transparency in this figure give rise to simultaneous perception of these two unique hues.

Colour constancy

Colour constancy is the ability of the visual system to keep the surface colour appearance of an object stable across changes in illumination (Hurlbert, 2007). This is a challenging task because, as stated in the previous section, light reflected from objects is a combination of the illumination and surface chromaticity. The brain must disentangle the illumination from the surface reflectance to create a stable representation of the object colour (Smithson, 2005). This is not a simple task, as the brain only receives information from three distinct cone receptors, and the illumination and surface reflectance signals are confounded. How does the brain overcome this challenge?

Colour constancy can be partially explained by adaptation of cone photoreceptor responses at the retinal level according to the coefficient rule of von Kries (von Kries, 1902). However, von Kries adaptation cannot account for all of our colour constancy ability, therefore there must also be higher level cortical processes involved. Some theories of colour constancy suggest that we use multiple cues from across the whole visual scene to calculate the illumination, for example by looking at the correlation between colour and luminance (Golz & MacLeod, 2002). The “gray world” hypothesis (Land, 1977; Land & McCann, 1971), suggests that the visual system calculates an achromatic point by averaging chromaticities across the visual scene. This helps identify the overall illumination of the scene, as the chromaticity of the achromatic point would be shifted colorimetrically towards the illumination.

Specular highlights from glossy surfaces could provide cues to the illumination which aid colour constancy (J. N. Yang & Maloney, 2001). As described in the previous section, a specular highlight on a glossy object primarily reflects the illuminant, whilst other regions of a surface reflect proportions of the surface reflectance chromaticity and

the illumination, in a gradient. When the chromaticities of a whole image containing the object are plotted (as in Figure 5.2), the plotted colours of the surfaces form a line that, when extrapolated, connects the object colour to the illumination colour. This is termed “chromaticity convergence “ (Hurlbert, 1998) and may aid in colour constancy calculations (H. C. Lee, 1986; Mollon, 2003).

Measuring colour constancy experimentally is challenging and researchers have approached this in different ways. Early colour matching experiments (Arend & Reeves, 1986) demonstrated that it is important to design experiments that distinguish between automatic adaptation processes underlying colour constancy, and more explicit inferential judgements (e.g. “if the light has changed, the appearance must have changed too”). It is also clear that the instructions given to observers affects their performance, as this may impact such inferential judgements (Radonjić & Brainard, 2016). A common way to measure colour constancy is to present two scenes under different illuminations (often rendered on a monitor). The participant is asked to make colour adjustments until the test patch under one illumination is a match the target surface under the other illumination, and to adjust so that it appears “cut from the same piece of paper”. This approach is known as asymmetric colour matching (Brainard, Brunt, & Speigle, 1997; Cornelissen & Brenner, 1995; Foster, 2011). However, there are limitations to this approach, as observers can simply learn the relationship between the test and target patch, and how they appear under different illumination, rather than genuinely creating a perceptual match. Another approach is ‘achromatic adjustment’ where observers adjust a surface in a scene until it appears white (containing no red, blue, yellow, or green) under various real or simulated illuminations (Brainard, 1998; Kraft & Brainard, 1999). It is clear that various mechanisms contribute to colour constancy, and it is achieved through a combination of low-level processes early in the

visual system, such as adaptation and lateral inhibition, and higher level processes such as extracting cues to illumination in order to factor illumination out of the light reflected from objects. However, it is still not yet fully understood how colour constancy is achieved, measuring it can be challenging and there has been doubts that it even exists (Foster, 2003; and see for reviews, Brainard, 2009; Brainard, Wandell, & Chichilnisky, 1993; Hurlbert, 1998; Smithson, 2005).

The development of colour constancy

We know that infants have fully functioning trichromacy by at least four months of age (Volbrecht & Werner, 1987). However, less is known about the development of higher level perceptual processes, such as colour constancy. Some theories of colour constancy posit that the visual system calculates illumination and surface reflectance by learning about the statistical regularities in our visual environment (Brainard et al., 2006).

However, these theories do not specify when such learning takes place. It is unclear whether we are born with implicit knowledge about the statistics of the natural environment (i.e. learned in our evolutionary history), or whether such learning occurs in early life.

The development of colour constancy appears to take much longer to fully mature, compared to colour discrimination abilities (Witzel, Sanchez-Walker, et al., 2013). This is an example of how the components of colour perception do not necessarily develop at the same rate. There have been numerous developmental studies on perceptual constancy, predominantly with infants using eye-tracking methods (e.g. Dannemiller, 1989; Dannemiller & Hanko, 1987; Granrud, 2006; Yang, Kanazawa, Yamaguchi, & Kuriki, 2013). These studies tend to focus on whether size constancy, lightness constancy, or colour constancy is at all present in young infants.

Evidence for perceptual constancy very early in life is informative, because it suggests that it is not an entirely learned capacity; we are born with rudimentary constancy abilities. However, there has been little investigation how rudimentary perceptual constancy in infancy develops into adult-like processes over the course of childhood. Furthermore, it has not been considered how immature perceptual constancy, and individual differences in maturation rate, may impact the development of other cognitive and linguistic domains.

1.5 Thesis overview and research questions

The current thesis investigates key components in the toolbox of colour perception, and draws comparisons between adulthood, childhood and infancy. The components of colour studied here are: colour constancy, colour term acquisition, the dimensions of colour, and surface and illumination discrimination.

The first research question asks whether the development of perceptual constancy in children impacts on other cognitive domains, such as language acquisition. Specifically, does the development of colour constancy in childhood constrain colour term acquisition? Previous work suggests there may be a link (Witzel, Sanchez-Walker, et al., 2013), but there had not been a highly controlled study to test the relationship. This question informs a broader debate about the relationship between perception and language. In paper 1, we studied first language acquisition and its relationship with perception.

Next, the dimensions of colour appearance were investigated using the psychophysical technique of Maximum Likelihood Conjoint Measurement (MLCM). In paper 2 and 3 we asked, how do the dimensions of colour appearance interact to inform judgements of colour in adults, and perceptual salience in infancy? Furthermore, do adults and infants differ in how the dimensions of colour appearance inform salience? The maturation of perceptual dimensions in infancy is also an important question in developmental psychology. Some have suggested that infants process stimuli in a more holistic way than adults, and dimensions become more separated with maturity (Kemler, 1983; Kemler & Smith, 1978, 1979; L. B. Smith & Kemler, 1978). In Paper 3 we demonstrated that MLCM can be combined with preferential looking methods to investigate interactions between dimensions of colour appearance in infancy. Finally,

we looked at surface change and illumination change discrimination, particularly in relation to colour constancy and the blue-yellow bias. Paper 4 asks why colour discrimination is poorest along the daylight locus.

Paper 1: Colour Constancy and Colour Term Acquisition are Positively Related in Early Childhood

Paper 1 investigated the relationship between colour constancy and colour term acquisition in three- to four-year-old children. We theorised that immature perceptual constancy may have an impact on the development of other cognitive domains, such as language acquisition. For example, if the colour of an object is unstable across changes in illumination due to undeveloped colour constancy, it may also be difficult to learn the colour term mapping to that object (e.g. yellow banana). Colour terms appear to be particularly difficult to learn for children, compared to other types of words (Bornstein, 1985b; Franklin, 2006; Soja, 1994). This can be particularly striking when, for example, a four-year-old may know all the complex names of her favourite dinosaurs, and yet mislabels a yellow ball as “green”. This has led some parents, including Charles Darwin, to wonder whether their child is colour blind (Darwin, 1877). There are many theories as to why colour terms appear to be particularly hard to learn. In paper 1, we investigated whether development of colour constancy may constrain colour term acquisition.

There is preliminary evidence for a relationship between colour term knowledge and maturity of categorical colour constancy in children (Witzel, Sanchez-Walker, et al., 2013). However, the method used in Witzel et al.’s study was limited by confounding colour term knowledge with the colour constancy task, thus making the relationship difficult to disentangle. Paper 1 tested the maturity of colour constancy in 42 children

aged three to four years, using a simultaneous matching task (Arend & Reeves, 1986).

In this task, participants are required to match a coloured target viewed under one illumination with one of four surfaces viewed under another illumination. The task was made into an age-appropriate game about finding matching coloured trousers for two bears (see Franklin, Clifford, Williamson, & Davies, 2005, for a similar task).

One stimulus was a “light match” to the target: it reflected the same wavelength of light despite being under a different illumination. Another stimulus was a “surface match” to the target: it was the identical stimulus to the target, and “cut from the same card” (colour constant choice). If a child had no colour constancy, they would mistake the “light match” stimulus for the target stimulus, as they would not be able to compensate for the shift in illumination. Two other stimuli were colorimetrically midway between the surface match and the light match. Stimulus choices were scored according to how colour constant the match was. In order to control for general ability in colour discrimination, the matching task was also conducted with no difference in illumination between stimuli. Colour term knowledge was assessed by testing production and comprehension of the eight basic chromatic colour terms.

Analyses revealed a positive correlation between colour constancy and colour term knowledge, even when taking into account the effect of age, and discrimination. These findings suggest that children who have more mature colour constancy also tend to know more colour words. Further research should now investigate the direction of this relationship, and investigate what other factors (e.g., sensitivity to the statistics of the visual environment or colour-object knowledge) are related to the maturity of colour constancy. This paper also demonstrates that colour discrimination develops more rapidly than colour constancy. This has implications for the findings in paper 4, where

we attempt to disentangle the effects of colour constancy on colour discrimination in adults.

Paper 2: Maximum Likelihood Conjoint Measurement of Lightness and Chroma

Paper 2 (Rogers, Knoblauch, & Franklin, 2016) uses the technique of Maximum Likelihood Conjoint Measurement (MLCM) to investigate how lightness and chroma interact in adults' perception of colour. MLCM is a psychophysical procedure which allows us to evaluate whether stimulus variation along one dimension influences perception along another (Knoblauch & Maloney, 2009). MLCM is based on a signal detection model of decision making that allows testing of several models of how observers integrate information to make choices. This technique is useful for studying colour perception, as colour is thought to be composed of three non-independent dimensions: hue, lightness and chroma (Wyszecki & Stiles, 2000). That is, a physical change in lightness (intensity) corresponds to a perceptual change in lightness, but also may cause a perceptual change in chroma. Previous literature on the interaction of hue, lightness and chroma in perception have been limited in their ability to quantify the relative weightings and interaction of the dimensions in perception (e.g. Burns & Shepp, 1988), whereas MLCM allows us to quantify this interaction.

In the first experiment, we created perceptual difference scales for lightness and chroma using Maximum Likelihood Difference Scaling (MLDS, Maloney & Yang, 2003).

Although the CIELCH colour space we used is designed to be perceptually uniform, it may be the case that the perceptual differences in these spaces are valid only over small differences, and are not reliable over larger distances (e.g., Hunt & Pointer, 2011; Wyszecki & Stiles, 2000). Adult observers were asked to judge the perceptual distance between stimuli that varied in either lightness or chroma, by indicating which of two

pairs of stimuli were perceptually lighter (or more chromatic). We used MLDS to analyse these judgements, and thus created perceptual difference scales for lightness and chroma.

In experiment 2, we used the scales developed in the first experiment to create matrices of blue, yellow, red and green stimuli that varied in both lightness and chroma and were perceptually equated in both dimensions. In separate blocks, observers judged which of two stimuli were a) lighter and b) more chromatic in a paired-comparison task. We then modelled how changes in one dimension influenced judgment of the other using MLCM.

We found that an additive model best fit the data in all conditions except for judgment of red chroma where there was a small but significant interaction. Lightness negatively contributed to perception of chroma for red, blue and green hues but not for yellow. We discuss various explanations for this, including that there was a linguistic component to the task. Observers were instructed to pick the “redder/bluer/greener/yellower” stimulus in the chroma judgement task, therefore they may have been influenced by the typical lightness of the focal colour of the term. The physical chroma of the stimuli tended not to influence observers’ perception of the lightness for yellow, green and blue colours. However, for red stimuli, a higher chroma increased the probability that the observer would judge it as “lighter”.

These findings suggest that the dimensions of colour are not independent: physical changes along one dimension can ‘contaminate’ a different perceptual dimension. Furthermore, this study demonstrates that MLCM can be used to study the interaction between the different dimensions of colour.

Paper 3: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy

Paper 3 combines the MLCM technique with the preferential looking method from developmental science, to investigate how dimensions interact to inform salience in infants. It is important to understand how physical dimensions, such as the dimensions of colour, interact in infancy. For example, the development of object recognition requires infants to process multiple dimensional properties of an object. Previous work has shown that the dimensions infants use to recognise objects (e.g. shape, size, pattern, colour) develops from 4 months to 12 months (Wilcox, 1999), and that covariation of multiple dimensional properties further aids in object identification (Woods & Wilcox, 2010). In adults, the dimensions of colour interact in perception (Burns & Shepp, 1988; Rogers et al., 2016). However, one preferential looking study found that 6-month-old infants tend to base their preference on hue, and luminance differences did not affect their preference (A. M. Brown & Lindsey, 2013). This suggests there may be developmental differences in how the dimensions of colour are weighted in perceptual judgements.

We tested twenty-two 6-month-old infants' and twelve adults' preferential looking to green stimuli that covaried in lightness and chroma and analysed preferences using MLCM. This improves on previous studies, as it allows the quantification and modelling of multiple dimensions' contribution to the decision process. With MLCM, we can quantify multi-dimensional contributions to perceptual salience. We define the most salient object in a scene as the one that is preferred, i.e. observers look first at it over its competitor in a forced-choice looking paradigm (Kaldy & Blaser, 2009).

The method involves eye-tracking observers' responses to pairs of stimuli, and coding which of the pair they look at first. For infants, the pairs of stimuli were randomly selected from a 3-by-3 matrix of green stimuli in which lightness and chroma independently vary. The analysis showed that infants' looking is driven primarily by lightness, with darker stimuli having a greater salience than lighter, plus a small but significant positive contribution of chroma. Whereas, adults' looking behaviour was mainly driven by chroma, with more chromatic stimuli being preferred over lighter stimuli. Both infants' and adults' looking "decisions" were based on an additive combination of the two dimensions of lightness and chroma. This indicates that infants do not separate the dimensions of colour to a greater extent than adults, as suggested by some previous studies.

However, there were differences between infants and adults in the way lightness and chroma contributed to salience. We discuss several explanations for this finding. First, infants may have been driven by contrast to the background. Second, adults' judgements may have been different since they were given instructions; and they may have been attempting to apply a strategy. Finally, 6-month old infants may be more sensitive to lightness than chroma due to greater magnocellular maturity than parvocellular maturity at this age (Dobkins, 2009). Magnocellular pathways are tuned to high temporal and low spatial frequencies and are more sensitive to luminance contrast, whereas the parvocellular pathways are more sensitive to red/green chromatic contrast (B. B. Lee, Pokorny, Martin, Valberg, & Smith, 1990; V. C. Smith, Pokorny, Davis, & Yeh, 1995). Previous studies have found that 3- and 4-month infants are more sensitive to luminance contrast than chromatic contrast (D. Allen, Banks, & Norcia, 1993; Dobkins, Anderson, & Lia, 1999). We hypothesise that differences between infants and adults found in our study may be linked to such pathway maturation. However,

parvocellular pathways can also underlie luminance change detection depending on the spatial structure of the stimulus, therefore further work would need to be done to test this hypothesis.

Paper 3 demonstrates that preferential looking behaviour can be used with a signal detection method to study the contributions of multiple physical dimensions to perceptual salience in development. This method could be extended to study the contribution of physical dimensions to other developmental processes such as face perception, language, and object recognition.

Paper 4: Why is Colour Discrimination Poorest along the Daylight Locus?

Paper 4 investigates why discrimination is poorest in the blue-yellow direction (roughly the same direction as the ‘daylight locus’) compared to the red-green direction in cone opponent colour space (Krauskopf & Gegenfurtner, 1992; MacLeod & Boynton, 1979). Several theories have been proposed to explain this asymmetric pattern in colour discrimination; some previous studies have suggested that it is functional to aid colour constancy (Pearce, Crichton, Mackiewicz, Finlayson, & Hurlbert, 2014; Radonjic, Pearce, et al., 2016). Colour constancy is the ability to discount the illumination colour from the surface colour (Foster, 2011), and most illumination in the natural environment falls along the blue-yellow line (Foster, Amano, & Nascimento, 2006). The “adaptive insensitivity” theory therefore posits that poor sensitivity to changes along the blue-yellow line would aid colour constancy, as the observer would not notice a change in natural illumination. An alternative theory proposes that insensitivity to colour changes along the blue-yellow axis is a by-product of colour constancy rather than an aid to it. This “constancy-noise hypothesis” suggests that uncertainty about the illumination adds noise to judgements along the blue-yellow axis (Bosten, Beer, & MacLeod, 2015).

However, there has not yet been a sufficiently controlled experiment in which discrimination of illumination change is compared to discrimination of surface colour change. If the theory that discrimination aids colour constancy is correct, it would be predicted that illumination discrimination would have a greater blue-yellow bias than surface discrimination.

In paper 4, we compared participants' surface discrimination ability to their illumination discrimination ability across a cone opponent colour space. We rendered highly controlled computer stimuli to accurately represent the real-world behaviour of light and surfaces. We created two sets of stimuli, which varied in a) the specularly reflected light (a cue to the illumination) or b) the diffusely reflected light (a cue to the surface reflectance), along eight different hue directions. We separately manipulated the specular and diffuse components of simulated reflected light in our stimuli to dissociate cues to illumination change from cues to surface reflectance change.

The experiment was a four-alternative forced-choice task in which observers determined which stimulus was the odd one out on each trial. Nineteen adult observers took part in both conditions. We found no significant difference between the conditions in the relative increase in discrimination thresholds along the blue-yellow colour axis (i.e. the daylight locus) compared to the orthogonal red-green axis. However, we did find a trend for relatively poorer discrimination in the blue-yellow direction for specularly reflected light than for diffusely reflected light. These findings do not provide strong support for the hypothesis that relatively poorer blue-yellow discrimination is adaptive for colour constancy, or a result of noise introduced by colour constancy. We suggest some alternative accounts, including the range-accuracy trade-off hypothesis, which suggests that the greater distribution of blue-yellow surfaces and illuminations in the natural

world leads to a coarser representation in the visual system. Further experiments could simulate a whole-scene illumination change, rather than a manipulating only the specularly reflected light and using this as a cue to illumination change.

1.6 Overall contribution

Contribution to understanding of the development of colour perception

Previous work has given us a good understanding of the maturation of the physiological structures supporting colour vision (Abramov et al., 1982; Yuodelis & Hendrickson, 1986), and how colour discrimination improves over development (Knoblauch et al., 2001). However, less is known from prior research about infants' perceptual response to colour, and the development of the components of colour vision such as the perceptual dimensions.

In this thesis, two papers further our understanding of the development of the components of colour vision. First, the findings of Paper 1 contribute to our understanding of colour constancy development, and second, Paper 4 informs us about how the dimensions of colour appearance interact to inform salience in infancy. Combined, these studies demonstrate that colour perception is still developing in infancy and childhood and not yet in its mature state. This parallels findings for the development of colour discrimination which also suggest that colour perception continues to mature into late childhood (Knoblauch et al., 2001). The different components of colour perception do not necessarily develop at the same rate, as these require different computations and representations which may have their own developmental trajectory. For example, components such as colour constancy require a representation of the surface chromaticity and the illumination.

Contribution to understanding the dimensions of colour appearance

The interaction of the dimensions of colour appearance in adults has been widely studied, as this is important to devising perceptually uniform colour spaces. However, there are limitations in the methods of prior studies, in that the relative contribution of

each dimension cannot be measured with methods such as multidimensional scaling (Ashby, 2007). Prior studies also left outstanding questions, for example about how context affects dimensional integration, and on these processes during infancy. Papers 2 and 3 used the method of Maximum Likelihood Conjoint Measurement to study interaction of the dimensions of colour appearance in infants and adults. This technique improves on previous work, as it allows the quantification of dimensions. The findings from these papers make a valuable contribution to the field, as they show that colour appearance dimensions interact in different ways over development, and this may also change across contexts. Paper 3 demonstrates that MLCM can be used in conjunction with preferential looking, using eye-movement as a proxy for judgement, enabling its use with developmental populations.

Contribution to understanding colour constancy and colour discrimination

Colour constancy is a large area of research, with many different experimental techniques and theories about how the brain achieves colour constancy (Foster, 2011; Smithson, 2005). Many theories of colour constancy propose that we use statistical regularities in the visual environment to help calculate illumination. One theory of colour constancy suggests that relatively poorer colour discrimination in the blue-yellow direction of cone opponent space is adaptive for colour constancy (Pearce et al., 2014; Radonjic, Pearce, et al., 2016). This theory points to a possible association between colour discrimination abilities and colour constancy. However, prior to this thesis the relationship between illumination discrimination measures and other measures of colour constancy has not been explored.

Paper 1 demonstrated that young children's failure to maintain colour constancy could not be explained by their poor colour discrimination. Furthermore, Paper 4 did not find

evidence to support theories that suggest that relatively poorer blue-yellow colour discrimination is adaptive for colour constancy. Together, these suggest findings that colour constancy in childhood and adulthood cannot be accounted for by colour discrimination abilities. The development of colour constancy is an under-researched area, but it is vital for our understanding of adult colour constancy mechanisms. The work in this thesis ties together colour constancy in childhood and adulthood, and demonstrates how both areas must be researched to understand the mechanisms.

1.7 Future research

There is exciting progress being made in our understanding of the components of colour perception, but there are still outstanding questions about their development and interaction. At the end of each paper in this thesis, there are suggestions of the next step to further develop each of the research questions tackled in this thesis. For example, the next step in investigating the relationship between colour constancy and colour term acquisition (Paper 1) is to conduct a longitudinal study. A longitudinal study would use similar methods to those used in Paper 1, but would test the same children at multiple time points between 36 months and 54 months of age. This would allow the investigation of colour term learning over time, and determine whether colour constancy maturity predicts rate of colour term acquisition and help disentangle the causal direction of the relationship. The next step to investigate the role of illumination perception in colour constancy (Paper 4) is to render full illumination change scenes, rather than changing only the specular component of the stimulus.

One major question is, what is the role of experience in colour perception and how are natural scene statistics learned? There is very little work on infants' sensitivity to natural scenes, or their sensitivity to deviations from statistical regularities in a scene. However, such work would be hugely informative to our understanding of when we learn about statistical regularities. For example, adult colour discrimination is poorest in the blue-yellow direction, and this appears to be related to the distribution of chromaticities in natural illumination (see Paper 4; Golz & MacLeod, 2002; Macleod & von der Twert, 2001; Pearce et al., 2014). Furthermore, there is evidence that colour perception is affected by season and latitude of birth, due to type of illumination exposure (Laeng et al., 2007). A future study could look at whether infants also show this bias in colour perception using preferential looking and discrimination thresholds.

Although work on perceptual plasticity tends to focus on infancy and childhood, it may also be the case that adults' visual systems are able to adjust to new input. For example, prior studies have used prism adaptation methods, in which participants wear goggles fitted with prism wedges that shift their visual field laterally or vertically. These studies show that adult participants successfully adjust their hand-eye coordination to the new visual input, even after very short amounts of time wearing the goggles (e.g. 3 minutes, Redding & Wallace, 1988). Furthermore, there is natural yellowing of the retina with age and the visual system can compensate for this without changing perception. This demonstrates that perception is remarkably plastic, even in adulthood.

It would be informative to study adaptation in adulthood to new chromatic scene statistics, and Virtual Reality (VR) techniques could be used in such experiments. VR techniques are starting to be recognised as useful for studying perception (Scarfe & Glennerster, 2015). Using a similar theoretical basis as the prism adaptation experiments, but with more sophisticated methods, adult participants would be asked to wear a VR headset where the scene statistics have been altered from the norm. For example, the distribution of surface chromaticities may be shifted from a blue-yellow favoured range to a red-green range. Participants' colour discrimination thresholds and colour constancy would then be tested. There is some prior evidence to suggest a change in the cue to the illumination alters colour constancy (Golz & MacLeod, 2002). If adults' colour perception can adapt to new scene statistics presented in VR, this would suggest that the visual system can become attuned to new environments even after childhood.

A second major research question is, how universal are the components of colour perception? Colour science tends to assume that the perceptual qualities of colour are universal. For example, the CIE (1931) perceptual colour spaces were based on

perceptual measurements from only 17 observers (Guild, 1932; Wright, 1929), and yet the colour space is used to produce stimuli for a wide range of people of various ages and backgrounds. Variation in the components of colour perception may be expected across individuals, cultures, or people from different environments based on several streams of prior evidence. First, there are differences in the world's colour lexicons (Roberson et al., 2000) and some have pointed to differences in the extent to which lexicons divide colour by lightness rather than hue (Maclaury et al., 1992). Second, whereas some have argued for universal trends in colour preferences (Franklin, Bevis, Ling, & Hurlbert, 2010; Hurlbert & Ling, 2007; Palmer & Schloss, 2010), a study comparing Himba to British adults found few similarities between the two groups (Taylor, Clifford, & Franklin, 2013). The Himba are a nomadic people living in North East Namibia and they tend to base their colour preferences on saturation, rather than hue as the British adults did. Both lines of investigation potentially challenge the idea that how the perceptual dimensions of colour contribute to colour perception is universal. Further cross-cultural investigations of the components of colour perception are important to test the idea that the components of colour perception are universal. Understanding any variation in the components could provide further leverage in identifying their underlying mechanisms.

The technique of MLCM used in Paper 3 and 4 would be a good way of investigating the universality of the dimensions of colour appearance. A cross-cultural study could be conducted using preferential looking in adult participants, comparing a non-urbanised population such as the Himba to an urban population such as British adults. The independence of the dimensions of colour in salience could be investigated. The hypothesis would be that the Himba population would order the dimensions of colour appearance in a different way to British population.

Creating an internal representation of the scene illumination supports colour constancy calculations (e.g. subtracting illumination from the surface reflectance). It may be the case that perceptual experience with scene illumination helps us to perform this calculation. There are differences between the chromaticities of artificial illumination in urban settings and natural illumination (i.e. sunlight and firelight), and therefore individuals living in different environments may differ in their colour constancy. How would colour constancy differ in individuals who mainly have exposure to natural, as opposed to artificial, light? A future study could compare individuals living in environments with primarily artificial illumination to those living in environments with natural illumination. This difference could be confirmed by asking participants to wear a device for a week that would record the chromaticities of their visual environment. The experiment would measure participants' colour constancy for changes in illumination along different hue angles, and compare the two groups. The hypothesis would be that participants exposed to artificial light would have better colour constancy for illumination changes not found in natural scene statistics (i.e. outside the daylight locus). Previous work lends support to the idea that there may be cultural and individual differences in the perception of the dimensions of colour appearance, and colour constancy. Cross-cultural experiments are important to challenge the idea that the components of colour perception are universal.

Finally, a third major question in the field is, where are the components of colour perception processed and represented in the brain? Papers 2 and 3 give evidence that stimulus variation along one dimension of colour can influence perception along another, and that the dimensions of colour appearance can interact to inform salience. There is evidence that the magnocellular pathway carries luminance information, and the parvocellular pathway carries chromatic information (B. B. Lee et al., 1990).

However, we know little about the neural processing of the dimensions of colour appearance. For example, where does luminance and chromaticity integrate? Future studies could use MLCM techniques in conjunction with neuroimaging, such as fMRI. Adult participants would make judgements (e.g. “which is lighter?”) about pairs of colours varying in luminance and chromaticity whilst in an fMRI machine. Analysis would reveal brain areas involved in making judgements associated with integration of the dimensions of colour, versus judgements associated with separation.

Preferential looking is often used in developmental studies to make conclusions about infants’ perception and attention (Teller, 1979); we use this technique in Paper 3. However, interpreting the results of preferential looking is challenging because it is difficult to know what is underlying looking behaviour. An eye movement may reflect, and is variously interpreted as: visual discrimination, preference, salience or attention (Aslin, 2007). To overcome limitations of preferential looking, and to study the location of cortical colour processing in development, future studies could use neuroimaging techniques. For example, electroencephalography (EEG) can be used with infants, and could be used to investigate changes in visual component associated with preferential looking to each dimension of colour appearance.

1.8 Conclusion

The following research illustrates that colour vision is composed of an array of components that vary in their availability and usefulness between contexts and life stages. At any one time, the brain may call upon one or more of the tools in colour perception to help it solve a given task. For example, young children may not yet have fully formed colour constancy, therefore they cannot call upon this tool to help them solve the task in Paper 1. Individual adults may differ in their ability to internally

represent a surface's illumination and chromaticity, as shown in Paper 4. Representation of the components of colour appearance changes over development, as shown in Paper 2 and 3, perhaps as a result of experience with the visual environment.

The work in this thesis demonstrates the components of colour perception, such as colour discrimination, colour constancy, and the dimensions of colour appearance and salience, work together to produce our day-to-day experience of colour. Future research should focus on colour perception in context, that is, how colour perception works for us in our daily lives, how individuals differ, and how task demands may impact performance. In sum, this thesis gives us a more complete understanding of contextual, developmental and individual differences in the components of colour, which serves to demonstrate that colour perception is not a monolithic process.

Chapter 2: Paper 1: Colour Constancy and Colour Term Knowledge are Positively Related in Early Childhood

Rogers, M. R.¹, Witzel, C.², Rhodes, P.³, & Franklin, A.¹

1 The Sussex Colour Group, School of Psychology, University of Sussex, UK

2 Laboratoire Psychologie de la Perception, Université Paris Descartes, Paris, France

3 School of Design, University of Leeds, UK

2.1 Abstract

The ability to keep perception constant despite environmental changes of illumination, viewing angle or distance is a key feature of human perception. Here, we investigate how ‘perceptual constancy’ relates to language acquisition by investigating the relationship between colour constancy and colour term knowledge in three to four-year-olds. We used a novel method to test colour constancy where children are required to match coloured stimuli under different illuminations. We found a positive relationship between colour constancy and colour term knowledge: children who knew more colour words also had better colour constancy. The relationship remained even when accounting for the effect of age and ability to discriminate colours. The findings have implications for understanding the development of perceptual constancy, language acquisition and the link between perceptual processing and cognitive development.

2.2 Introduction

We can recognise an apple under many different types of illumination and at varying viewing angles and distances, even though the image reaching the retina varies greatly between those contexts. The brain extracts a stable representation from varying input, giving us a backdrop to detect the important changes (Walsh & Kulikowski, 1998). This process, known as perceptual constancy, also enables us to label, and therefore communicate, about objects independent of contextual variation (e.g. “apple”). In childhood, we are faced with the challenge of acquiring words and applying them consistently, and a stable perceptual environment is likely to be key in helping children with the challenge of language acquisition (Garrigan & Kellman, 2008).

Colour is an elementary feature of our visual environment. Some argue that children find it more difficult to learn colour terms than other types of words (e.g. Bornstein, 1985; Soja, 1994), although the rate of colour-term acquisition also varies greatly between children (Kowalski & Zimiles, 2006). It is unclear why colour, which is a seemingly simple perceptual attribute, would be a challenge for children to name. One potential factor is that colour naming requires dividing up millions of discriminable colours into relatively few categories. For example, there are many variations of colour that would all be termed “green”. To acquire colour terms, children must learn how to map terms onto these colour categories and identify the boundaries of colour categories in their language: where “green” becomes “blue” (Wagner et al., 2013).

A further challenge is that the colour of a surface varies due to illumination changes. The light reflected from the surface of an object changes depending on the illumination of the object, leading to different tristimulus colour signals when the light reaches the retina. The ability to identify the colours of objects and surfaces despite changes in

illumination is called colour constancy (Foster, 2011; Smithson, 2005). For example, if you are drinking from a white cup and you move from outside (illuminated by the sun), to indoors (illuminated by incandescent light bulbs), the light reflected from that cup would change. However, our brain discounts the illumination and keeps a constant perceptual experience of a white cup. The adult visual system usually achieves colour constancy remarkably well (Abrams, Hillis, & Brainard, 2007). Some of the effects of illumination are already compensated at an early, sensory stage of colour processing, such as adaptation or local contrast (Hansen, Walter, & Gegenfurtner, 2007; Kraft & Brainard, 1999). Therefore, it is unsurprising that evidence for rudimentary colour constancy, adaptation and local contrast are found to be already present during infancy (Dannemiller, 1989; Dannemiller & Hanks, 1987; Pereverzeva & Teller, 2009; J. Yang et al., 2013). However, low-level sensory mechanisms cannot fully explain colour constancy (Foster, 2011; Smithson, 2005), therefore higher-level processes must be involved in disentangling changes of colour that are inherent to the object, from changes of colour that are due to the illumination. For example, scene interpretation plays an important role in colour constancy (Mollon et al., 2017; Winkler, Spillmann, Werner, & Webster, 2015; Witzel et al., 2017).

Additionally, if we consider all theoretically possible surfaces and illuminations, there are infinite possible shifts in the colour of a surface under illumination changes (Logvinenko, 2009; Logvinenko & Tokunaga, 2011), but very few of them occur in our visual environment (Foster, Amano, & Nascimento, 2006). Hence, colour constancy may also require the observer to learn which colour shifts can be expected in the visual environment (Witzel, van Alphen, Godau, & O'Regan, 2016). To be able to reliably name the colours of objects, a child has to distinguish both the variability of colours within and across colour categories (e.g., Wagner, Dobkins & Barner, 2013), and the

variability of colours due to objects and illumination changes. Therefore, the challenge of colour constancy may contribute to the difficulty of colour-term acquisition. If so, children with advanced colour constancy could also be advanced in colour naming as a result. Whilst good colour constancy could support colour term acquisition in this way, colour term acquisition could also aid colour constancy: having terms for the colours of objects may help anchor perception under environmental change. The first step in distinguishing between these theoretically important possibilities, is to establish whether the development of colour constancy and the acquisition of colour terms are related.

In adults, colours that are named most consistently across observers are also named most consistently across illumination changes, demonstrating a link between colour naming and colour constancy changes (Olkkonen, Hansen, & Gegenfurtner, 2009; Olkkonen, Witzel, Hansen, & Gegenfurtner, 2010). Three- to four-year-olds who can group colours well with colour terms, are also better at keeping these colour groupings constant under different illuminations (Witzel, Sanchez-Walker, et al., 2013). However, these studies are concerned with the consistency or constancy of naming and grouping colours across illuminations. They do not demonstrate that someone who is better able to name colours is also better able to keep a colour constant under changes in illumination. The current study aims to establish whether colour constancy and colour naming are related in 3-4 year old children who are still learning colour terms. Are children with more advanced colour naming also better at keeping colours perceptually constant across illumination changes?

To measure colour constancy in the current study, children are required to match a physical coloured target under one illumination to one of four surfaces viewed under a different illumination. The task is made into a developmentally-appropriate game about

matching the colours of trousers for two bears (adapted from Franklin, Clifford, Williamson, & Davies, 2005). One stimulus was created to be a “light match” to the target: it reflected light with the same sensory colour signal (CIE xyY ; u^*v^*) as the target when under a different illumination. Another stimulus was the “colour constant match”: it consisted of the same surface as the target and cut from the same coloured card. The other two stimuli were colorimetrically midway between the light match and the colour constant match. This means that one stimulus under red light has the same colour measurements as the other stimulus under white light.

A child without colour constancy would see the light match stimulus as identical to the target under the two different illuminations. On the other hand, a child with adult-like colour constancy would not be ‘fooled’ by this, and would recognise the stimuli cut from the same piece of paper under the two lighting conditions. The midpoint stimuli allow us to test graduations of colour constancy. Each trial was scored from 0 to 3: the higher the score, the closer the child’s choices were to the colour constant match. The experimental setup is such that the child would not adapt fully to one illumination or the other, as they are simultaneously visible, thus likely resulting in adaptation to a midpoint. However, this does not present a problem to the method as this would make the task more challenging and prevent a ceiling effect in the results. To control for general task ability such as children’s ability to match and discriminate colours, the matching task was also conducted with no difference in illumination between stimuli (i.e. the task required only colour discrimination). Colour-term knowledge was assessed by testing production and comprehension of the eight basic chromatic colour terms. If the development of colour constancy and colour-term acquisition are linked, we predict that children who are better able to match colours correctly despite changes in illumination will also be better at naming and comprehending colour terms. Such a

Chapter 2: Colour Constancy and Colour Term Knowledge are Positively Related in Early Childhood

finding would provide the first clear evidence for a link between the development of perceptual constancy and language acquisition.

2.3 Methods

Participants

Forty-nine children (21 girls) with no family history of colour vision deficiency took part. Seven children were excluded due to completion of less than 50% of the trials (mean age = 34 months, SD = 3.5 months). Data from the remaining 42 children (17 girls) were retained for analysis. They were between 33 months and 45 months of age (mean age = 37 months, SD = 3.1), 86% went to nursery (full time or part time) and they had an average birth weight of 3.34kg (SD = 0.51). All children were presented with a small gift (a book) at the end of the study as a thank you for participating. This study received ethical approval from the Science and Technology Cross-Schools Ethics Committee (C-REC) at the University of Sussex and the European Research Council Executive Agency Ethics Committee (ref 230685).

Colour constancy task

Apparatus and stimuli

Two stimuli were custom printed using a Hewlett-Packard Designjet Z3200 large-format printer on HP Premium Matte photo paper. This printer uses a 12-ink system to achieve a wide colour gamut and its pigmented inks are claimed to be lightfast for 200 years (“Datasheet: HP DesignJet Z3200 Photo Printer series,” 2017). The stimuli were matched to target values precisely ($0.3 \Delta E_{00}$) and were designed such that card 1, “*purple*”, under D65 illumination (which simulates daylight) was a metameric light match of card 2, “*blue*”, under red Lee filtered light. This means that when card 2 is measured under red light, it is the same colour as card 1 under natural light: they are a “light match”. This was confirmed by measuring the card under the two illuminations with a SpectraScan PR-655 spectroradiometer (Photo Research Inc., Chatsworth, CA; see Table 2.1 and Figure 2.1 for measurements).

Two additional stimuli were then selected in matte Munsell card to fall roughly between the measured chromaticity and luminance values (CIE $u'v'$) of the two custom printed cards (see Figure 2.1). CIE u' and v' are the red-green and blue-yellow axes of a perceptual colour space. The additional midpoint stimuli had a Munsell value of 5 PB 5/6 and 7.5 PB 4/6, we have labelled these as “*blue-purple*” (*b-p*) and “*purple-blue*” (*p-b*), respectively. The midpoint stimuli were indistinguishable from the printed test card in all ways except for the colour (e.g., identical in texture), therefore type of card could not be a cue in the task. The training stimuli were matte Munsell card, in clearly discriminable colours, but with a similar luminance (Munsell values: 2.5B 7/8; 7.5B 3/6; 5Y 8/12; 5R 5/12).

All stimuli were cut into trouser shapes, of the same dimensions (9cm x 4.5cm). We also created white cardboard cut-outs of bears for the task (18cm x 9.5cm). We used a customised lightbox in this task to control the illumination. The box was separated into two compartments: one side was lit by unfiltered D65 simulator bulb (VeriVide fluorescent D65, length 600mm, wattage 18, diameter 38mm) and the other side had a red Lee filter (number 035) covering the aperture resulting in red-filtered light from the same source. We refer to these as “white illumination” and “red illumination” respectively. See Figure 2.2 for photos of the setup.

Table 2.1. The CIE xyY and $u'v'$ values of the test stimuli under white illumination and red filtered illumination. The rows in bold show that the purple stimulus under white light is a metameric match of the blue stimulus under red light.

	<u>White illumination</u>					<u>Red filtered illumination</u>				
	<u>x</u>	<u>y</u>	<u>Y</u>	<u>u'</u>	<u>v'</u>	<u>x</u>	<u>y</u>	<u>Y</u>	<u>u'</u>	<u>v'</u>
<i>blue</i>	0.242	0.266	26.56	0.170	0.419	0.261	0.244	15.33	0.193	0.403
<i>blue-purple</i> (<i>b-p</i>)	0.246	0.253	22.09	0.177	0.410	0.266	0.234	12.98	0.202	0.399
<i>purple-blue</i> (<i>p-b</i>)	0.242	0.227	18.86	0.185	0.390	0.263	0.213	8.95	0.210	0.381
<i>purple</i>	0.261	0.236	14.95	0.197	0.400	0.285	0.223	7.39	0.223	0.393

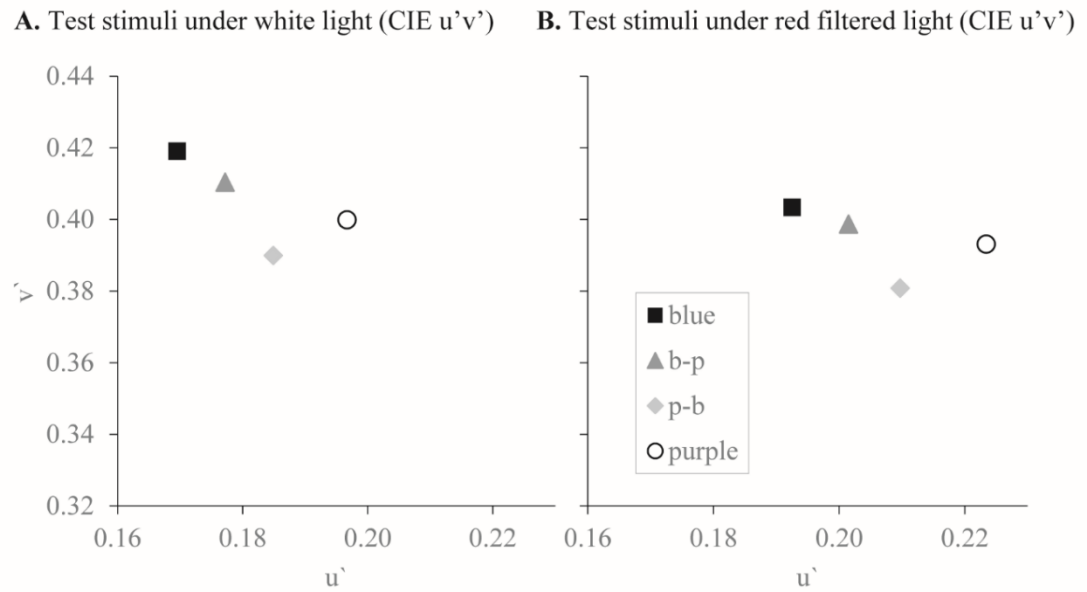


Figure 2.1. Colorimetric measurements of the four test stimuli measured in CIE $u'v'$ under A) white light (D65) and B) red filtered light. These figures show that b-p and p-b fall roughly midway between the blue and purple stimuli in CIE $u'v'$, a perceptual colour space. The coordinates of the purple stimulus (unfilled circle) under white light, are the same as the blue stimulus (black square) under red light.

Procedure

Colour constancy was measured with a simultaneous matching task, where target and response stimuli had different illumination. We also conducted control trials with all stimuli under the same illumination. The child sat in a chair facing the lightbox, Experimenter 1 sat next to the child, the child's parent or carer sat close by, and Experimenter 2 sat behind them. Experimenter 1 read the child a story book introducing the concept of matching coloured trousers on bears. The story consisted of photos of the bear cut-outs and stimuli used in the experiment, and showed two bears wearing the same coloured trousers set in different everyday scenes (e.g., at the park).

Experimenter 1 then turned on the lightbox whilst Experimenter 2 simultaneously turned off the main light illuminating the room. Thus, the room was completely dark apart from the light from the lightbox. Experimenter 1 explained the task whilst the child adapted to the illumination. The training trials then began. Experimenter 1 put the cardboard cut-out bears (see Figure 2.2) in the white light section of the lightbox with the four training stimuli the child could choose from. Experimenter 1 also held four of the same colour stimuli hidden in her hand. The child was reminded that these two bears always liked to wear the same colour, then placed one stimulus from her hand on a bear. "Remember that these bears always like to wear the same trousers. Can you make them match?" Experimenter 1 placed one trouser-shaped stimulus from her hand onto the bear and asked the child to place a stimulus from the four options in the lightbox on the other bear to make them match. When the child had made a match, the procedure was repeated. The stimuli placement was shuffled between trials. The participants did not receive any feedback on the selection. The training phase finished when the child got three in a row correct.

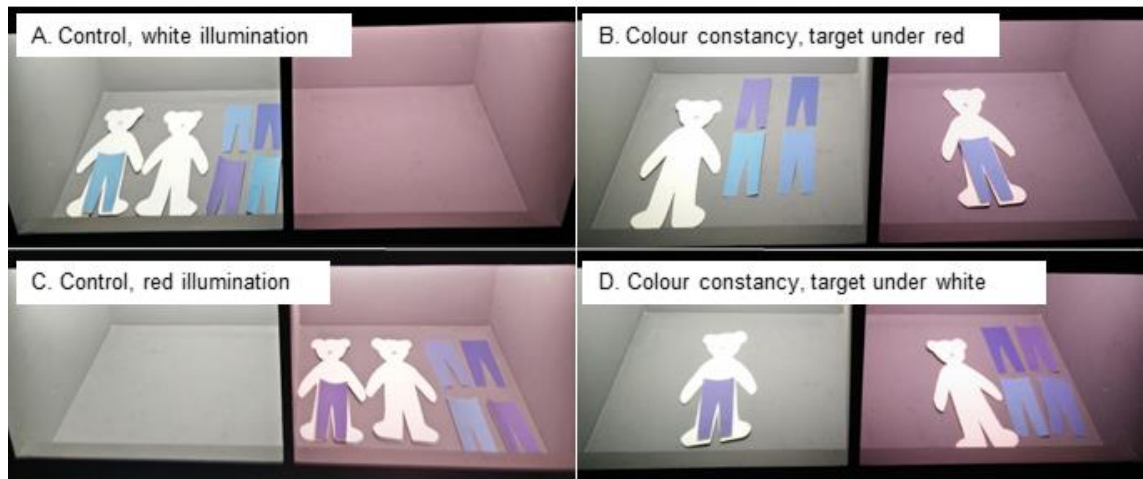


Figure 2.2. Representation of the layout of the lightbox and the position of the stimuli in the four test conditions: A) Match target stimulus to options under white illumination B) Match target stimulus under red illumination to options under white illumination; C) Match target stimulus to options under red illumination and D) Match target stimulus under white illumination to options under red illumination.

For the experimental phase, Experimenter 1 replaced the training stimuli with the four experimental stimuli in the lightbox (see Figure 2.2). The procedure was the same as the training trials. In the control trials, all stimuli were presented under the same illumination (i.e. all stimuli were in one compartment of the box).

Experimenter 1 placed a target stimulus on one bear, and then the child was asked to make them match by selecting the corresponding stimulus from the four options laid out. For the purple colour constancy trials, the purple target and one of the bears were placed in the white illuminated side of the box, whilst the other bear and four options were placed in the red illuminated side. For the blue colour constancy trials, the blue target and one of the bears were placed in the red illuminated side of the box, whilst the other stimuli were in the white illuminated side. The target was only ever the ‘blue’ or ‘purple’ stimulus in the discrimination and colour constancy trials. Experimenter 2 noted down the child’s selection.

The red filter was counterbalanced between left and right position across participants. The task order was either A, B, C, D or C, D, A, B, so that trials alternated between control and colour constancy (see Table 2.2 and Figure 2.2 for trial types). Each type of trial was conducted three times. The trial order was not fully counterbalanced as this would require moving the selection stimuli between compartments for each trial. A pilot study showed that children lost attention and completed fewer trials when the stimuli had to be moved between each trial.

Table 2.2. Four different types of trial were conducted, that assessed the child's ability to match the target stimuli to the selection stimuli, under the same or different illumination.

<u>Trial type</u>	<u>Task</u>	<u>Illumination for selection stimuli</u>	<u>Target stimulus</u>	<u>Illumination for target stimulus</u>
A	Control (match under same illumination)	Neutral illumination	Blue	Neutral illumination
B	Colour constancy	Neutral illumination	Blue	Red illumination
C	Control (match under same illumination)	Red illumination	Purple	Red illumination
D	Colour constancy	Red illumination	Purple	Neutral illumination

Coding

For each trial in the colour constancy task, the child's target stimulus selection was appointed a score from 0 to 3. When the target was *purple*, zero points were given for a selection of *blue*, one point for a selection of *blue-purple*, two points for a selection of *purple-blue* and three points for a selection *purple* (i.e. target match). When the target was *blue*, the awarding of points were reversed. See Table 2.3 for a full list of scores.

If the child chooses a target colorimetrically further away from the target under the same illumination, it indicates that the ability to discriminate that colour from the target is low; hence, the further the chosen colour, the lower the discrimination performance. By calculating points for each response, this allowed us to calculate a graded colour constancy score and colour discrimination score.

An overall *colour constancy score* and a *discrimination score* were calculated for each participant, by summing their points for each trial (i.e. a number out of 3) within a condition, and then calculating this as a percentage of the highest number of points possible for the number of trials they attempted. The equation to calculate the colour constancy score is:

$$\text{Colour constancy score} = \frac{\sum \text{Scores from all CC trials}}{\text{Highest score possible from CC trials attempted}} \times 100$$

Equation 2.1

The discrimination score calculation is the same, except using the control trial scores. Not all children completed the whole set of trials; 88.76% of all possible trials were completed. Scores were calculated out of number of trials that the child attempted rather than number of all possible trials. This measurement best represents children's comprehension, and reflects their true colour constancy capacity rather than their fatigue

Chapter 2: Colour Constancy and Colour Term Knowledge are Positively Related in Early Childhood

with the task. The following analyses find the same results if measures are calculated out of all possible trials.

Table 2.3. Points for colour constancy and colour discrimination tasks

<u>Selection</u>	<u>Points when target blue</u>	<u>Points when target purple</u>
blue	3	0
blue-purple (b-p)	2	1
purple-blue (p-b)	1	2
purple	0	3

Colour naming task

Stimuli

For the colour term comprehension task, we used an image of a rabbit surrounded by eight different coloured clothes on laminated card. These colours were chosen to be good examples of the eight basic chromatic colour terms: blue, green, red, yellow, pink, purple, orange and brown. For the colour term production task, we used images of a rabbit wearing each of the different coloured clothes on separate pieces of laminated card (as in Pitchford & Mullen, 2002).

Procedure

Following the colour constancy task, the main light was switched back on, and the child was presented with the card described above. The child was asked to point to the colour as Experimenter 1 said the colour term (“Where is the *red* jumper?”). This was done for each colour term. For the colour term production task, the child was shown each of the eight colour naming cards and asked to say the name of the colours (“What colour jumper is the rabbit wearing here?”); the colour naming cards were shuffled between participants.

2.4 Results

The average number of colour terms participants could comprehend was 6.90 out of 8 ($SD = 1.56$), and the mean number of colour terms participants could produce was 6.69 out of 8 ($SD = 1.51$). Pearson's correlation showed that colour term comprehension and production were positively correlated ($r = 0.705, p < .0001$). Therefore, to investigate the relationship between colour naming and colour constancy, we averaged together colour term comprehension and production to produce a colour naming score for each participant.

The colour constancy scores were non-normally distributed, as indicated by the Shapiro-Wilk Test ($p = .004$) with skewness of $-.581$ ($SE = 0.365$) and kurtosis of -0.818 ($SE = 0.717$). Therefore, Spearman's rank correlations were carried out on the data. Spearman's rank correlation indicated a positive relationship between colour naming and colour constancy ($\rho = .521, p < .001$). Further, more stringent, correlations were carried out using the robust correlation toolbox (Pernet, Wilcox, & Rousselet, 2013), to ensure outliers do not contribute to the correlation. The bend correlation and the skipped correlation protect against outliers, and we still find a relationship using these measures (bend correlation: $r = 0.500, p < .001$; skipped correlation: $r = 0.521, p < .05$). This suggests that children who know more colour terms have a higher level of colour constancy. See Figure 2.3 for a scatterplot showing the relationship between colour naming and colour constancy score. Piloting revealed that adult participants achieved 100% scores on this task.

We also calculated a "colour constancy ratio", in which we divided the raw colour discrimination score by the raw colour constancy score. Thus, individual discrimination abilities are used as a baseline and taken into account. A ratio of 1 indicates equal

performance on both conditions, and the lower the ratio the poorer colour constancy score is relative to the colour discrimination score. The colour constancy ratios were also non-normally distributed according to the Shapiro-Wilk Test ($p = .004$; skewness = $-.437$, $SE = 0.365$; kurtosis = -0.822 , $SE = 0.717$). Spearman's rank correlation indicated a positive relationship between colour naming and colour constancy ratio ($\rho = .553$, $p < .001$).

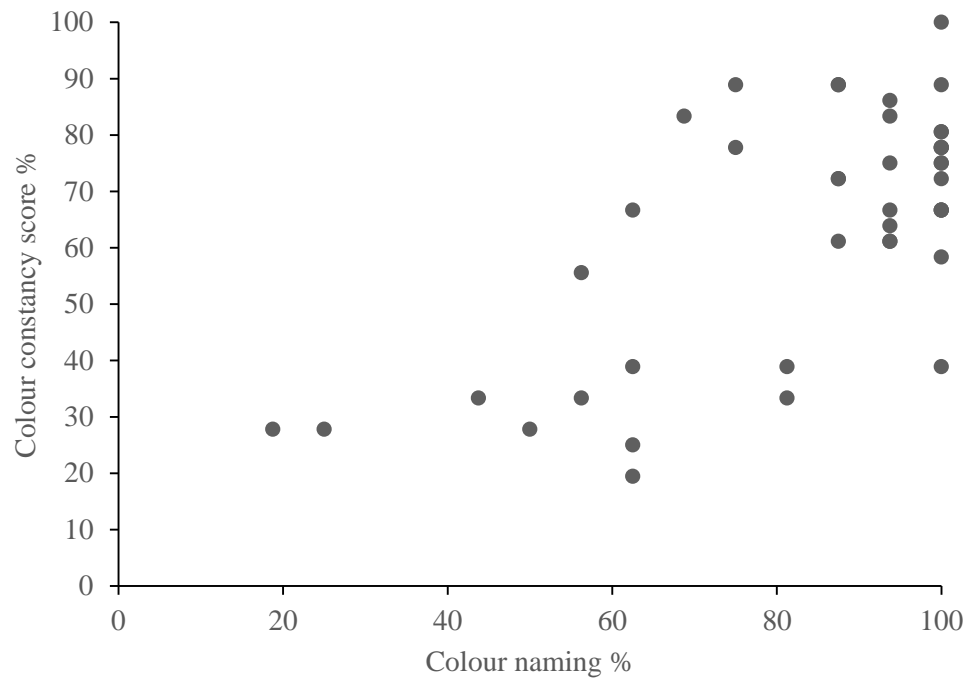


Figure 2.3. Colour constancy score and colour naming ability for each participant.

Colour naming is calculated by averaging the colour term comprehension and colour term production scores, converted to a percentage. There is a positive relationship between the two variables. $N = 42$.

When colour discrimination is controlled for in a partial Spearman's correlation, the relationship between colour constancy and colour term knowledge remains ($\rho = 0.549, p < .001$), which suggests that poor performance on the discrimination trials cannot account for the relationship between colour constancy and colour term knowledge. Furthermore, the relationship remains when controlling for both age and colour discrimination score ($\rho = 0.587, p < .001$). This suggests that it is not simply the case that older children are better at both colour naming and the colour constancy task, thus driving the relationship between colour naming and colour constancy score. See Figure 2.4 for a scatterplot of colour naming and colour discrimination score. There was also no relationship between colour naming and age ($\rho = 0.085, p = .591$), suggesting that age is not the driving factor explaining individual differences in colour naming.

We also conducted Bayesian analyses on the relationships of interest. Bayes factors allow interpretation of the strength of the evidence for either the null or alternative hypothesis (Dienes, 2014). A Bayes Factor (B) of 0.33 or lower indicates evidence for the null hypothesis, and a B of 3 or above indicates substantial support for the alternative hypothesis. Values between 0.33 and 3 (closer to 1) suggest that the data is not sensitive enough to concretely support either the null or alternative hypothesis (Wetzels & Wagenmakers, 2012). As there has been little prior work in this field, and due to the non-parametric nature of the data, it is difficult to predict priors (i.e. the pattern of data we expect). Therefore we conducted Bayesian non-parametric correlations (Kendall's Tau) and used the default beta width prior of 1, which assigns equal prior probability to all correlation values between -1 and 1 (van Doorn, Ly, Marsman, & Wagenmakers, 2016). These analyses provided support for the experimental hypothesis in the relationship between colour naming and colour

constancy score, $\tau = 0.390$, $BF_{10} = 126.8$; support for the null hypothesis in the relationship between colour naming and discrimination score, $\tau = 0.073$ $BF_{10} = 0.251$ and support for the null hypothesis in the relationship between age and colour naming $\tau = 0.072$ $BF_{10} = 0.249$. The data were not sensitive for the relationship between age and colour constancy score, $\tau = -0.185$ $BF_{10} = 0.857$.

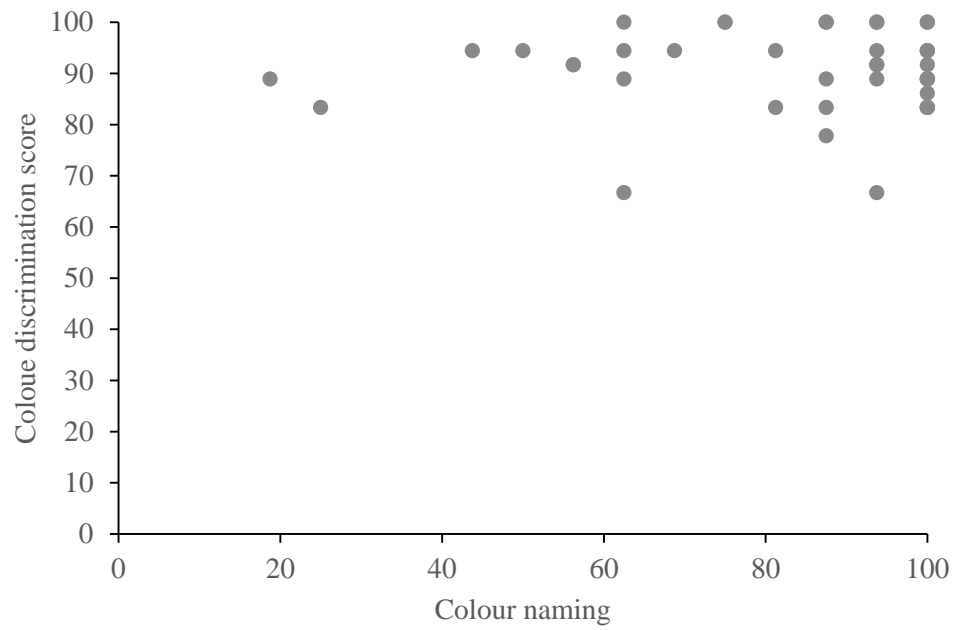


Figure 2.4. Colour discrimination score and colour naming ability for each participant.

There is no relationship between the two variables. $N = 42$.

2.5 Discussion

This study investigated the relationship between colour constancy maturity and colour-term knowledge, in three to four year olds. We found a positive relationship between colour-term knowledge and colour constancy. We found no relationship between colour-term knowledge and colour discrimination, nor between age and colour constancy. We can be confident that the relationship between colour constancy and colour-term acquisition is not due to children's ability to understand the task, general task demands, or their age, since children's colour-term knowledge does not relate to their ability to do the task when there is no illumination change and older children were not more colour constant. These findings suggest that children who have more mature colour constancy also tend to know more colour words and vice versa. This may be because immature colour constancy makes it challenging to learn colour terms by mapping them to coloured objects (e.g., yellow banana), as the colour of the object would change with an illumination shift. Furthermore, colour-term knowledge may help to 'anchor' the representation of colour during illuminant changes, via top-down influence of colour naming (Witzel, Maule, et al., 2013; Witzel, Sanchez-Walker, et al., 2013)(Witzel, Maule, et al., 2013; Witzel, Sanchez-Walker, et al., 2013)(Witzel, Maule, et al., 2013; Witzel, Sanchez-Walker, et al., 2013).

This study demonstrates that two to four year olds do not yet have fully mature colour constancy, and that individual variation in constancy may impact the development of other cognitive domains, such as language acquisition. Previous studies have found that colours that are named most consistently across observers are also named most consistently across illumination changes, when participants were asked to sort colours into categories under different illuminations in adults (Olkkonen et al., 2009, 2010), and in three to four year olds (Witzel, Sanchez-Walker, et al., 2013). Our findings provide

the first clear support for a relationship between perceptual constancy and language acquisition, when constancy is measured with a task that is independent of naming.

In line with previous literature on colour-term acquisition (Bornstein, 1985b; Franklin, 2006; Kowalski & Zimiles, 2006; Soja, 1994), our findings demonstrate that even at three to four years old, children are still acquiring mature colour-term usage. In the current study, some children could name 100% of the eight basic chromatic colour terms whereas, some children could name fewer than 20%. We did not find a correlation between colour-term knowledge and age, which suggests that other factors are driving the individual differences in colour-term knowledge. This finding may seem surprising given that children acquire language as they develop, and therefore it might be expected that older children have greater colour-term knowledge than younger children.

However, the absence of a correlation between age and colour-term knowledge likely reflects the wide range of individual differences in rate of colour-term acquisition, as found in previous studies (e.g. Kowalski & Zimiles, 2006).

Prior studies have revealed rudimentary constancy abilities in young infants using preferential looking techniques (Dannemiller, 1989; Dannemiller & Hanko, 1987; J. Yang et al., 2013). However, with these methods it is difficult to interpret whether the infant is showing a preference for the novel or familiar stimulus (Houston-Price & Nakai, 2004), which means that inferences about constancy can be difficult to make. Furthermore, some of these studies used monitor-rendered simulations of illumination changes rather than real-world stimuli, and it is unclear whether infants would be able to interpret these simulations as an illumination change. In the current study, we adapted a colour constancy task from the adult literature (i.e. simultaneous matching; Arend & Reeves, 1986) to be suitable for young children, and we used real-world illumination

changes rather than monitor-based simulations. This task can now be used to further chart the development of colour constancy and to understand the processes that underlie the development of this important perceptual skill.

The equipment restraints in the experiment meant that the control trials had a different physical setup to the colour constancy trials. That is, the colour constancy trials were conducted across two sections of the lightbox whereas the control trials were conducted in one section. This may have meant that the colour constancy task was more challenging than the illumination task, beyond the change in illumination challenge. However, this was the same across all participants and thus would not affect the finding that children with greater colour term knowledge had higher colour constancy.

Perceptual constancy is a key feature of our visual system, and investigation into constancy is informative for engineering solutions, such as in computer vision (Vazquez-Corral, Vanrell, Baldrich, & Tous, 2012) and machine learning (Agarwal, Gribok, & Abidi, 2007; Gouko & Kobayashi, 2010). Therefore, further work into the development of constancy would be highly informative for many areas. We know that there is individual variability in the constancy abilities of adults (E. C. Allen, Beilock, & Shevell, 2011, 2012), further research into individual differences during development could shed light on what makes a “constant brain”. Future work could also adopt methodological innovations from the adult constancy literature. For example, adult colour constancy studies have used tuneable LED lighting systems to illuminate a room and test constancy over changes in illumination (Pearce, 2015; Radonjic, Aston, et al., 2016). This is far more realistic than screen-based simulations of illumination change, yet still precisely controlled. This method could be used adapted into a realistic, child-

appropriate measure of colour constancy, for example a task involving finding and retrieving objects in the illuminated room.

In conclusion, there is a positive relationship between the maturity of colour constancy and colour-term knowledge in young children. This demonstrate the impact of perceptual constancy on the maturation of other cognitive domains during childhood, such as language acquisition. It also raises questions about the development of colour constancy through children's interaction with their visual environment.

Chapter 3: Paper 2: Maximum Likelihood Conjoint Measurement of Lightness and Chroma

Rogers, M., Knoblauch, K., & Franklin, A. (2016). Maximum Likelihood Conjoint Measurement of Lightness and Chroma. *Journal of the Optical Society of America A*. 30(3), A184-A193. doi: 10.1364/JOSAA.33.00A184.

3.1 Abstract

Colour appearance varies along dimensions of lightness, hue and chroma. We used Maximum Likelihood Conjoint Measurement to investigate how lightness and chroma influence judgments of colour appearance. Observers judged lightness and chroma of stimuli that varied in both dimensions in a paired-comparison task. We modelled how changes in one dimension influenced judgment of the other. An additive model best fit the data in all conditions except for judgment of red chroma where there was a small but significant interaction. Lightness negatively contributed to perception of chroma for red, blue and green hues but not for yellow. The method permits quantification of lightness and chroma contributions to colour appearance.

3.2 Introduction

The dimensions of colour appearance

Colour appearance can be specified in a three dimensional perceptual space, composed of the dimensions *lightness*, *hue* and *saturation* (or *chroma*). Perceptual colour spaces, such as those defined by the CIE or the Munsell colour system, attempt to characterize lights and surfaces according to dimensions, so that the space approximates perceptual uniformity (Munsell, 1912; T. Smith & Guild, 1932). In this paper, we seek to evaluate whether stimulus variation along one dimension influences perception along another. The results of such an investigation may be complex, for example, if the underlying psychological dimensions prove to interact non-additively and lack independence. When this occurs, the dimensions are called ‘integral’ as opposed to ‘separable’ (Goldstone, 1994). In colour, such interaction has been demonstrated for similarity judgments in which observers’ classifications of stimuli varying in lightness and chroma were found to depend upon overall similarity rather than that predicted by their dimensional components (Burns & Shepp, 1988). It would be expected, then that stimulus dimensions that are ‘integral’ cannot be processed separately, or disentangled in perception (Garner, 1974; Lockhead, 1972), even when integration of information from multiple dimensions hinders task performance (Saarela & Landy, 2015).

Several methods have been exploited to investigate the dimensional structure of colour perception. For example, the method of direct estimation (Cohen, 2003; Newhall, Nickerson, & Judd, 1943), requires that observers explicitly choose stimuli of equal spacing along a continuum. This is a highly subjective task, and observers show considerable variability when the experiment is repeated. It is also prone to task-unrelated bias, such as that the difference between a given pair of stimuli influences

judgments of a subsequent pair (Krantz, 1972). Multidimensional scaling has also been used to investigate the interaction of colour dimensions in perception (e.g. Bimler, 2011; Burns & Shepp, 1988; Ekman, 1954; Helm, 1964; Indow, 1988; Indow & Kanazawa, 1960). This technique uses similarity or difference judgments or ratings between stimuli to determine the number of dimensions necessary to describe the perceptual space of the stimuli and the distribution of stimuli in that space. However, it does not specify what the actual dimensions are (Krantz, 1972).

Given the limitations of previous methods, there is interest in techniques that can model how multiple perceptual dimensions interact to constrain observers' judgments. It would be particularly useful to be able to quantify individual differences as there is evidence that some groups of people may weight colour appearance dimensions differently to others. For example, an analysis of 6-month old infants' looking time data from four previous studies led the authors to hypothesize that pre-linguistic infants' colour preferences ignore lightness, and only pay attention to hue (A. M. Brown & Lindsey, 2013). There is also evidence that experience with colour appearance dimensions may influence observers' ability to separate dimensions. For example, colour scientists are better than non-experts at extracting information about one dimension of a colour, without being influenced by the other dimensions (Burns & Shepp, 1988). Furthermore, participants trained to categorize a new hue-based colour boundary (not pre-existing in the basic colour term lexicon) also extended their learning to lightness-based distinctions. This may reflect the integrality of these two dimensions, i.e. attending to hue variation in the presence of (task unrelated) lightness variation resulted in sensitization to lightness distinctions (Burns & Shepp, 1988).

Conjoint measurement

Conjoint measurement is a psychophysical method designed to investigate how specific dimensions contribute to perception (Falmagne, 2002; Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971; Luce & Tukey, 1964). It is based on paired-comparisons of stimuli covarying independently along two or more dimensions. The conjoint-measurement approach is useful because it allows the investigation of more than one variable, the construction of measurement scales and allows us to determine how the variables combine in perception (Krantz & Tversky, 1971).

Ho, Landy and Maloney (2008) recently recast the decision process of the conjoint measurement task within a signal detection framework which because of explicit assumptions about noise in the judgment process, permitted the perceptual scales underlying the judgments to be estimated by a maximum likelihood procedure, yielding Maximum Likelihood Conjoint Measurement (MLCM). Ho *et al.* (2008) showed that the data could be fit with a series of three nested models, corresponding to three different decision rules of how observers combine the responses of the dimensions: (i) the independence model, in which the judgments depended on only one of the dimensions, (ii) the additive model, in which the judgments depended on an additive combination of component response functions from each dimension and (iii) the saturated model, in which an interaction term was needed to model the judgments beyond the additive contributions of each dimension. An added benefit of this approach is that the estimated scales can be specified in terms of the signal detection parameter d' (Knoblauch & Maloney, 2012). The technique has been used to investigate mutual influences of surface roughness and glossiness in texture perception (Ho *et al.*, 2008; Qi, Chantler, Siebert, & Dong, 2015) and to measure how the strength of the Watercolor Effect is

influenced by several stimulus dimensions (Gerardin, Devinck, Dojat, & Knoblauch, 2014).

Using virtually the same psychophysical procedure, Kingdom, Bell, Haddad and Bartsch (2015) examined the influence of chromatic and luminance components on perceived blur. However, their analysis was quite different, in that they did not estimate the underlying perceptual scales directly, but inferred independence based on the estimated probabilities of choosing a stimulus as stronger (blurrier in their case), when the physical component was in fact stronger. Another approach for characterizing how psychological dimensions combine is based on General Recognition Theory (Soto, Vucovich, Musgrave, & Ashby, 2015). This method, also, uses a signal detection approach to quantify the different contributions of dimensions to judgments, based on a stimulus identification task. One limitation of this method is that the task and model can become very complicated if there are many levels and stimuli.

The current study

In the current study we used MLCM to investigate how chroma and lightness contribute to chroma and lightness judgments for several hues. In order to maximize the possibility of interactions between these two dimensions and to simplify eventual interpretations, we sought to use stimuli that extended over a similar perceptual range and that were equally spaced perceptually along each dimension. To accomplish this, we conducted a preliminary experiment using Maximum Likelihood Difference Scaling (MLDS, Knoblauch & Maloney, 2008, 2012; Maloney & Yang, 2003). MLDS is a psychophysical procedure used to estimate perceptual scales for stimuli distributed along a single physical continuum. The technique requires observers to make judgments comparing the perceptual intervals between pairs of stimuli (e.g., between which pair of

stimuli is the difference greatest?). The signal detection model and maximum likelihood estimation procedure yield a scale with interval properties, i.e., equal scale intervals are perceptually equal. Since the task requires observers to report which interval between pairs of stimuli is greater and not simply which stimulus is stronger, as in discrimination experiments using paired-comparisons, it allows larger stimulus difference to be evaluated. MLDS has previously been used in colour research to equate colour differences perceptually, specifically to ensure equivalence of perceptual differences between targets and distractors for a visual search task (Lindsey et al., 2010). It has also been used as an independent measure of the perceived difference between stimuli (A. M. Brown, Lindsey, & Guckes, 2011), as so-called uniform colour spaces are notoriously non-uniform (Hunt & Pointer, 2011; Wyszecki & Stiles, 2000). Interestingly, it was found that perceived differences estimated by MLDS better predicted visual search performance than colour category boundaries, contrary to previous work without this control.

In the second experiment we used MLCM (Ho et al., 2008; Knoblauch & Maloney, 2012) to quantify the ‘contamination’ of lightness when making judgments about chroma, and vice versa. In separate experiments, participants were asked which of two stimuli (varying independently in lightness and chroma) were (i) lighter or (ii) more chromatic. The model assumes that the human visual system calculates a perceived lightness (or chroma) for the stimuli, which could depend on responses to both their physical lightness *and* chroma, and that the observer uses this information to make a judgment. Here, lightness and chroma are described as “physical dimensions”, despite the fact that units in colour space are often described as “perceptual”. This is to distinguish between the manipulations of the stimuli by the experimenter in the physical world, and the internal psychological/perceptual response of the observer. Experiments

were conducted for four hues (red, yellow, blue and green), to determine whether there are different patterns of lightness and chroma interaction for different hues. We found that lightness and chroma interact in an additive manner when making judgments about the lightness or chroma of a pair of stimuli.

3.3 Experiment 1: Methods

We designed the first experiment to choose stimuli to be used in the subsequent experiment. We used Maximum Likelihood Difference Scaling (MLDS, Knoblauch & Maloney, 2008; Maloney & Yang, 2003) to estimate four perceptually equidistant levels along the dimensions of chroma (when hue was set to red, yellow, blue or green) and achromatic lightness, spanning a perceptually equal range across dimensions.

Observers

Ten observers participated in the experiment (female = 7; mean age = 25.3, SD = 2.83, range = 23 – 32). Five observers completed lightness, red chroma and yellow chroma judgments. An additional five observers completed green chroma and blue chroma judgments. All were assessed as having normal colour vision using Ishihara plates (Ishihara, 2010).

Apparatus

Stimuli were presented on a 22 inch 2070SB Mitsubishi Diamond Plus Diamondtron CRT monitor, with a resolution of 1600 x 1200 pixels, 24 bit colour resolution and a refresh rate of 100Hz. Experiments were performed in a dark room. The experiment was run using custom software written in MATLAB R2012b and the Psychophysics Toolbox extensions. The monitor was calibrated using a ColorCAL colorimeter (Cambridge Research Systems).

Stimuli

Stimuli were specified in CIE LCh_{uv} , as chroma and lightness can be varied independently in this space. CIE LCh_{uv} is a cylindrical version of CIELUV, where L is the lightness, C^*_{uv} is the chroma, and h_{uv} is the hue (Poynton, 2012). Ten levels of five

dimensions were selected: red, yellow, blue and green chroma and achromatic lightness. Chroma is the colourfulness of a colour judged against the brightness of another colour that appears white under similar viewing conditions, as opposed to saturation which is the colourfulness of a surface relative to its own brightness. Colourfulness is the degree of difference of a colour from achromatic grey (Commission Internationale de L'Eclairage (CIE), 2017). Hues angles were selected on the basis of the category exemplars for English speakers (Sturges & Whitfield, 1995). For the lightness condition, chroma = 0 and hue = 0. In the chroma condition, red = 14.3° , yellow = 80.2° , blue = 234.3° and green = 143.2° hue angle in LCH space. Lightness was fixed at 50 for the chroma condition. Red chroma had a higher range because the monitor gamut was much wider at the specified lightness level ($L = 50$), whereas for yellow, blue and green, higher chroma values were not possible within the gamut. A grey background (xyY (1931): 0.31271, 0.32902, 50) was used throughout the experiment, and this was also used as the whitepoint for colourimetric conversions. A whitepoint is important to specify the simulated illumination the stimuli are under, and to allow the stimuli to be reproduced across different monitors.

Table 3.1. CIE LCh_{uv} Values for the 10 Levels of Stimuli in Experiment 1. The whitepoint for obtaining these values was xyY (1931): 0.31271, 0.32902, 50

Level	Lightness	Chroma (C in LCH space)			
	(L in LCH space)	Red	Yellow	Blue	Green
1	25	11	5	5	5
2	30.5	21	10	10	10
3	36	31	15	15	15
4	41.5	41	20	20	20
5	47	51	25	25	25
6	52.5	61	30	30	30
7	58	71	35	35	35
8	63.5	81	40	40	40
9	69	91	45	45	45
10	74.5	101	50	50	50

Table 3.2. xyY values for the 10 levels of stimuli in the MLDS experiment. Y for all chromatic hue conditions at all levels was 9.21, and x y for all levels of the achromatic lightness condition was 0.313, 0.329.

	Red		Yellow		Blue		Green		Achromatic lightness
	x	y	x	y	x	y	x	y	Y
1	0.337	0.330	0.321	0.341	0.302	0.321	0.309	0.339	2.208
2	0.358	0.331	0.330	0.354	0.291	0.312	0.305	0.349	3.221
3	0.380	0.332	0.339	0.367	0.281	0.304	0.301	0.360	4.504
4	0.401	0.333	0.349	0.381	0.270	0.296	0.297	0.371	6.090
5	0.422	0.334	0.359	0.395	0.261	0.289	0.292	0.382	8.010
6	0.443	0.335	0.369	0.410	0.251	0.281	0.288	0.394	10.296
7	0.464	0.336	0.380	0.426	0.241	0.274	0.283	0.407	12.981
8	0.484	0.337	0.392	0.442	0.232	0.267	0.278	0.420	16.095
9	0.504	0.338	0.404	0.459	0.223	0.260	0.273	0.434	19.672
10	0.524	0.339	0.416	0.477	0.214	0.253	0.267	0.449	23.743

Design and Procedure

On each trial of the experiment, three stimuli were presented, horizontally in a row on the monitor (a “triad”). Observers were asked to judge whether the left or right stimulus was more similar to the middle stimulus. The stimuli varied along one of the following five dimensions: lightness, red chroma, yellow chroma, blue chroma or green chroma. The values along each dimensions were chosen from a set of ten pre-determined values indicate in Table 3.1.

The triads were always ordered in intensity from left to right on the screen, randomly assigned to ascending or descending order along the dimension (e.g. either levels 1 2 3 or 3 2 1; but never 1 3 2). Every element in a triad differed in stimulus level, that is, there were no repeats of the same level within a triad (e.g., no 1 2 2). For 10 levels along each dimension, there are 120 unique triads.

A trial began with presentation of a central fixation cross for 200ms, followed by a triad presented until the observer responded, thereby initiating the next trial. Observers were instructed to judge whether the left or right stimulus was more similar to the middle stimulus. They responded by pressing the left or right button on a button box.

Observers’ responses were coded as their choice of left or right (0/1). One group of observers ($N = 5$) judged red, yellow and achromatic triads, resulting in 360 trials. A second different group of observers ($N = 5$) judged blue and green triads, resulting in 240 trials. Trials were randomized, interleaved and presented in one session with blocks to allow for breaks.

3.4 Experiment 1: Model

The analysis is described in detail elsewhere (Knoblauch & Maloney, 2012). A difference scale was estimated from each session using functions from the MLDS package (Knoblauch & Maloney, 2008) in the open source software R (R Core Team, 2017), and the scales from individual observers were averaged to obtain means and standard errors for the estimated scale values.

The analysis is derived from a signal detection model in which the observer's judgments depend on a decision variable based on comparing perceptual intervals between pairs of stimuli, e.g. with stimulus triad (a, b, c) , intervals (a, b) and (b, c) . It is assumed that on each trial the decision variable is perturbed by Gaussian noise with variance σ^2 .

Intuitively, if for a given stimulus b , the observer is equally likely to choose either of the two stimulus intervals; the perceptual intervals between the two pairs are equal.

Scale values and the variance are estimated by maximum likelihood so as to predict best the set of the observer's responses over the course of an experiment. The estimated scale has the property that equal scale differences correspond to equal perceived differences. The scale is unique, however, only up to a linear transformation, i.e., adding and/or multiplying all scale values by a constant does not affect the predictions. Thus, we fix the scales to be zero at the lowest stimulus values tested and to have $\sigma^2 = 1$ at each stimulus level. Parameterized in this fashion, the estimated scale values of an MLDS experiment are on the same scale as the measure d' from Signal Detection Theory (Green & Swets, 1966; Macmillan & Creelman, 2005), because when the standard deviations of the two distributions are equal sensitivity may be represented by d' (Abdi, 2009). That is, in units of the standard deviation of the internal noise (see Devinck & Knoblauch, 2012). We used this parameterization in the graphs throughout the article.

3.5 Experiment 1: Results

Figure 3.1 shows the average perceptual scale values across observers within conditions. The red chroma condition displays a steep slope at low chroma values indicating a “crispening effect” (Spehar & Zaidi, 2000; Takasaki, 1966, 1967; Whittle, 1992), i.e., enhanced sensitivity to chroma differences at low chroma values. Subsequent to this high slope region, the scale values increase approximately linearly. The scales for the other dimensions are similar in their trends and more nearly linear though the yellow and blue scales display a slight crispening at low chroma and the slope of the achromatic scale decreases slightly at high values.

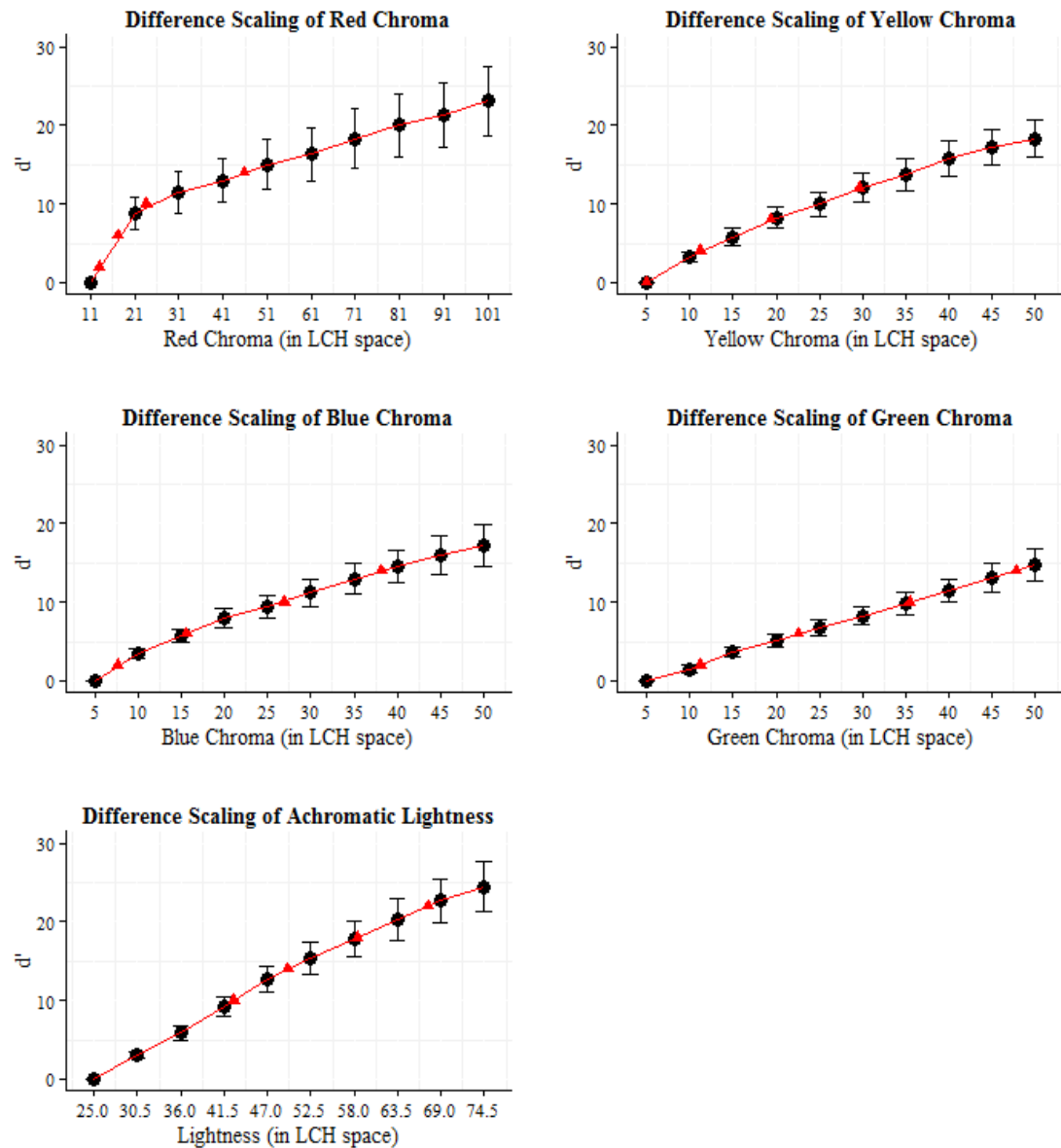


Figure 3.1. Average perceptual scales for each condition. The black dots show mean data, and the (red) triangles show four interpolated points, each separated by 4 d' (although not all starting at the same place). The error bars show ± 1 SEM. Refer to online version for colour.

Interpolation

The purpose of conducting MLDS was to choose a stimulus set spanning perceptually equal ranges across dimensions with perceptually equidistant levels along each dimension for use in the subsequent MLCM experiment. Four levels along each dimension were selected with a range of $d' = 12$, the maximum value within the gamut along all dimensions and then interpolating values in LCH space (on the x axis) from the specified d' (y axis) values. These points are indicated on the graphs in Figure 3.1 as red triangles and their LCH_{uv} values are shown in Table 3.3. The units for the chroma conditions for each hue are specified in C and the achromatic lightness condition in L^* .

Table 3.3. CIE LChuv values for interpolated points ranging in d' from 0-12

<u>Dimension</u>	<u>d' scale</u>			
	<u>0</u>	<u>4</u>	<u>8</u>	<u>12</u>
Chroma Red	12.96	17.19	23.66	45.79
Chroma Yellow	5.00	11.25	19.40	29.60
Chroma Green	11.26	22.59	35.47	47.79
Chroma Blue	7.65	15.60	26.99	38.19
Lightness	42.75	49.55	58.47	67.33

3.6 Experiment 2: Methods

As outlined in the general introduction, Maximum Likelihood Conjoint Measurement is a psychophysical procedure that models how two physical dimensions contribute to a perceptual judgment (Ho et al., 2008; Knoblauch & Maloney, 2012). In this experiment, the two investigated dimensions were Lightness and Chroma. The levels of these dimensions were selected to be perceptually equidistant using the MLDS procedure outlined in experiment 1.

Observers

Thirty observers participated in the experiment (female = 21; mean age = 20.8, SD = 3.05). All were tested for colour vision deficiencies using the Ishihara plates (Ishihara, 2010). Observers were paid £8 per hour for their participation.

Apparatus

The same apparatus was used as in experiment 1. A new program was written in MATLAB R2012 using Psychophysics Toolbox extensions to run the experiment (Brainard, 1997; Kleiner, Brainard, & Peilli, 2007; Pelli, 1997).

Stimuli

Four levels of Lightness and four levels of Chroma that were identified in Experiment 1 as giving four equally spaced points perceptually were used for each hue. These formed a 4 x 4 matrix of stimuli varying along the dimensions of lightness and chroma for each of the hues: red, yellow, blue and green (see Figure 3.1 for an illustration of the green stimulus matrix).

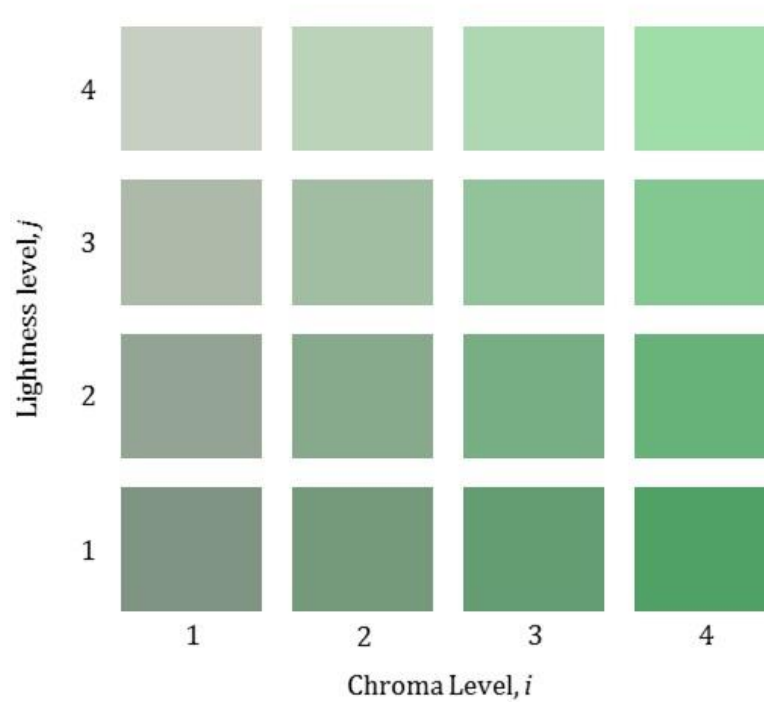


Figure 3.2 Representation of a stimulus set for green hue, varying along four lightness levels and four chroma levels. Note that they will not appear correctly in print or on an uncalibrated monitor. The first stimulus in a pair is specified by a chroma level (i) and a lightness level (j), the second stimulus is also defined by chroma level (k) and lightness level (l).

Design and Procedure

On each trial, a pair of stimuli chosen at random from the 4 x 4 grid (Figure 3.2) were presented and observers were asked to judge which one was lighter (condition 1) or which was more chromatic (condition 2). Given the 16 pairs of stimuli per hue condition, there were $16 * 16/2 = 128$ pairs to test (self-comparisons were included but not analysed as they do not contribute to the estimated scale values). Each pairs was presented twice with left-right positions reversed, resulting in 256 pairs. This procedure was repeated twice, resulting in a total of 512 trials per condition. Each condition took 10-15 minutes, resulting in a 40-60 minute experiment.

A trial consisted of a central fixation cross presented for 200ms, followed by a pair of stimuli presented on screen until the observer made a response (left or right) with the button box, which initiated the next trial. Observers were asked “Which stimulus is lighter” or “Which stimulus is redder/yellower/bluer/greener” (depending on the hue condition). Two sets of observers participated; the first judged red and yellow (N=15), and the second judged blue and green hue stimuli (N=15). Observers saw both sets of hue stimuli twice, once judging lightness and once judging chroma. This results in 4 experiments per observer (2 hues and 2 judgments) which were counterbalanced in a Latin square design.

3.7 Experiment 2: Model

It is assumed that when the judging the lightness of stimuli i,j and k,l , the observer forms the noise-contaminated decision variable:

$$\Delta(i, j, k, l) = \Psi^L(\varphi_k^l, \varphi_l^c) - \Psi^L(\varphi_i^l, \varphi_j^c) + \varepsilon.$$

Equation 3.1

According to the model, if the decision variable (Δ) is negative, then the observer chooses the left stimulus, whereas if it is positive, she chooses the right stimulus. The Ψ^L terms are interval responses to lightness that is a function of the stimulus lightness and chroma. Similar terms (Ψ^C) appear for chroma when the observer judges chroma. The noise term, ε , is included in order to account for the fact that observers will not necessarily make the same choice on repeated trials when stimulus differences are small. Given a specific form for the combination of responses from the stimulus attributes, the MLCM model estimates the scale values $(\varphi_k^l, \varphi_l^c)$ by maximum likelihood so that the estimated scale values best predict the observers' choices over the experiment.

We considered three nested models to fit the data, with the first the most constrained and the last the least constrained. First, the independence model assumes that the observers' judgments depend on only one of the component dimensions. For example, there is no contamination of chroma (lightness) when making lightness (chroma) judgments; only the difference in lightness (chroma) between the two stimuli influences the observers' choices. Second, the additive model assumes that the observer's response to a stimulus is a simple sum of the component psychological responses to the physical dimensions. Third, the saturated model has the highest number of free parameters and

allows for an interaction between the dimensions. The model is termed saturated because it includes the maximum number of parameters to estimate given the number of stimulus pairs presented.

Initially, we fit the additive model to the data using the MLCM package (Knoblauch & Maloney, 2014). This model assumes that the physical lightness level (ϕ^l) and chroma level (ϕ^c) of the stimulus separately and additively contribute to perceived lightness (Ψ^L) and perceived chroma (Ψ^C). Here, upper case L and C are used to refer to perceived dimensions, whereas lower case l and c are used to refer to physical dimensions. For the additive model, the decision variable is estimated by:

$$\Delta(i, j, k, l) = [\Psi_k^{L:l} - \Psi_i^{L:l}] + [\Psi_l^{L:c} - \Psi_j^{L:c}] + \varepsilon,$$

Equation 3.2

where $\Psi_k^{L:l}$ is an additive contribution of physical lightness to perceived lightness that is constant along a row (i) in Figure 3.2 and $\Psi_k^{L:c}$ is the contribution of physical chroma and is constant along a column. The additive model for lightness judgments estimates the difference between the perceived lightness of the two stimuli plus an additive contribution of perceived chroma. The independence model is obtained by suppressing the responses to one of the dimensions in the equation above. The saturated model is obtained by including an interaction term that depends on the levels, i and j of each stimulus, shown in the equation below:

$$\Delta(i, j, k, l) = [\Psi_k^{L:c} + \Psi_l^{L:l} + \Psi_{kl}^{L:cL:l}] - [\Psi_k^{L:c} + \Psi_l^{L:l} + \Psi_{kl}^{L:cL:l}] + \varepsilon.$$

Equation 3.3

3.8 Experiment 2: Results

The additive model was fit to judgments of lightness and chroma for the four conditions of red, yellow, green and blue, to estimate how much each dimension contributed to the decision using the MLCM package (Knoblauch & Maloney, 2014). The average contributions of chroma and lightness to the judgments for each stimulus condition are shown in Figure 3.3. The column labels indicate the judgment made by the observers and the row labels indicate the hue of the chromatic component of the stimuli. Circles indicate the lightness contribution and triangles the corresponding chromatic contribution to the judgments. The ordinate values indicate the 4 stimulus levels for each component as indices varying from 1 – 4. This corresponds to the 4 levels of lightness and chroma used in the experiment. The additive model assumes no interaction between the levels, therefore the data can be plotted along just four levels (rather than 4 x 4 levels).

In Figure 3.3, the top left plot shows the additive model fit to judgments of the redness of stimuli. Interestingly, lightness negatively contributes to judgments of the redness of stimuli. This means that a higher lightness component in a stimulus will tend to diminish its chromatic appearance leading the observer to be less likely to choose that stimulus as more chromatic than a stimulus of equal chroma but lower lightness.

Chroma of the stimulus positively and linearly contributes to judgments of redness. A similar behaviour was observed for both the blue and green stimuli, but notably not for the yellow. For the yellow stimuli, increased lightness slightly increases the chromatic appearance of the stimuli.

The second row right plot shows judgments of lightness for yellow stimuli. Here, there appears to be little contribution of the chromaticness of the stimuli to judgments,

because the chroma component curve is close to zero at all stimulus intensity levels. The lightness component however positively contributes. The same relations are observed with the blue and green stimuli but not for the red. For the red, the data show a small contribution of chroma to lightness for the most chromatic stimuli. Recall that we limited the gamut of the stimulus sets so that the range of stimuli would be perceptually equal. This seems to be born out in that the range of responses for the judged component are approximately equal across all conditions. The contributions of the levels of the dimensions to judgments also looks to be linear in the additive model graphs. The linearity of the responses likely reflects the pre-selection of equally spaced stimuli with MLDS method.

The additive model corresponds to just one possibility to describe the data. Another possibility is that the judgments depend on the contributions of only one of the dimensions, as appears to be approximately the case for the chromatic contributions to lightness for three of the hues. This model has been referred to as the “independence” model and is nested within the additive model (Ho et al., 2008; Knoblauch & Maloney, 2012). There is also a more complex model in which the additive combination of the responses does not suffice to describe the data and an interaction term is included to describe the deviations from additivity of the component responses. This model is referred to as the “saturated” model because it includes the maximum number of parameters to describe all of the stimulus conditions.

The additive and independence model are nested within the saturated model (Knoblauch & Maloney, 2012). We could fit and evaluate the three models with a likelihood ratio test for each individual but this would entail performing multiple tests (30 observers

with 2 tests per observer would give 60 tests) which would neither be efficient computationally nor statistically.

In the software used to fit the models (Knoblauch & Maloney, 2012), the maximum likelihood procedure is implemented as a generalized linear model (GLM) with a binomial family (McCullagh & Nelder, 1989). This suggests the more efficient alternative of estimating the values using a generalized linear mixed-effects model (GLMM) in which average (population) responses are estimated as fixed-effects and observer sources of variability are estimated as random effects (Bates, Maechler, Bolker, & Walker, 2014; Knoblauch & Maloney, 2012; Moscatelli, Mezzetti, & Lacquaniti, 2012). In this, responses across observers are combined, which may confound potential observer differences in criterion and sensitivity. A GLMM, however accounts for the multiple sources of variance both within and between observers. Each observer takes part in four experimental conditions. There is variation in performance between conditions because there are different instructions. Furthermore, different observers are likely to use different criteria in their judgement. Variance between observers or between multiple sessions by the same observer are called random effects. The fixed effects represent the estimates of the population averages.

To fit such a model, we take advantage of the fact that because of our choice that stimuli be equally spaced perceptually, the estimated scales are approximately linear as a function of the stimulus indices, and additionally, they pass through the origin by construction. Thus, to simplify the GLMM model specification, we will treat each response scale as a linear function of the indices. Then, we only need to estimate the slopes for each component scale. In this case, the fixed effects component of the decision variable for the additive model can be notated as:

$$\Delta_{ijkl} = (\beta_L L_i + \beta_C C_j) - (\beta_L L_k + \beta_C C_l) + \epsilon,$$

Equation 3.4

where the β 's are the slopes (or gains) and L and C the indices of the luminance and chroma components, respectively and ϵ is the Gaussian distributed judgment noise of the observer. This can be rearranged and simplified as

$$\Delta_{ijkl} = \beta_L (L_i - L_k) + \beta_C (C_j - C_l) + \epsilon = \beta_L \Delta L_{ik} + \beta_C \Delta C_{jl} + \epsilon.$$

Equation 3.5

Thus, the decision variable is a function of only the differences between stimulus indices. Any offsets or intercepts in this formulation would be cancelled by the differencing of levels. In the independence model, we constrain one of the slopes to be zero and for the saturated model, we introduce a coefficient β_{CL} applied to a product of the indices to obtain an additional interaction term. The advantage of the GLMM formulation, however, is that we can also estimate and test variance components associated with individual differences in the slopes of the components for each condition. For example, a full GLMM model including both fixed and random effects can be notated as:

$$g(E[Y = 1]) = (\beta_{L, TX} + b_{L, O} + b_{L, OT} + b_{L, OX} + b_{L, OTX}) \Delta L_{ik} + (\beta_{C, TX} + b_{C, O} + b_{C, OT} + b_{C, OX} + b_{C, OTX}) \Delta C_{jl}.$$

Equation 3.6

In this model, the dependent variable, Y , codes the choices of the observer (left/right as 0/1) and their expected value, here, is the probability of choosing the right (the

probability of left being 1 minus this value). This expected response is related to the linear predictor through a link function, g , that we take to be an inverse cumulative Gaussian, making this a type of probit analysis. The linear predictor on the right is the sum of the luminance and chromatic difference terms, each with a complex sum of components determining its gain. Greek letters are used for the fixed effect terms. The fixed-effect term indicates an average or population estimate across the data set. The subscripts L or C indicate whether the term corresponds to a luminance or chroma slope, T , (indexing lightness or chroma tasks) and X , the hue of the stimulus set. Roman letters are used for the random effects whose variance contributions are estimated. The term with subscripts L , O accounts for observer specific variations in the slopes for the lightness component with a similar term for the chroma component. Successive random effects, $b_{L,OT}$, $b_{L,OX}$ and $b_{L,OXT}$ refer to random observer effects of the slope of the luminance component that are specific to the task (T), the hue component (X) or both the task and the hue component with matching terms for the chroma slope. Not all of these random effects need be required to model the data, and we evaluate them by performing likelihood ratio tests on nested models fit with and without specific random effects, to establish the simplest variance structure that best describes the data sets. After having established the variance structure, we then continue by testing the fixed-effects with respect to the three nested models described above: independence, additive and saturated. The GLMM models were fit using the function *glmer* function in the **lme4** package (Bates et al., 2014) from the OpenSource software R (R Core Team, 2017). The data from the two groups of observers (red/yellow and blue/green experiments) were fit separately.

The initial evaluations of the random effects indicated that the observer and observer, condition variance components were not significant for the luminance term in the red and yellow experiments. For the blue and green experiment, the Observer variance for the luminance term was also not significant but the Observer, hue Condition term associated with the chromatic term ($b_{c,ox}$) was significant. Likelihood tests for the fixed effects in the red and yellow experiments rejected the independence model ($\chi^2(4) = 74.9, p \ll 0.001$) and the additive model ($\chi^2(4) = 12.3, p = 0.015$), thus supporting an interaction in the combination of lightness and chroma for these conditions. On the other hand, for the blue and green experiments, the independence model was rejected ($\chi^2(4) = 55.7, p \ll 0.001$) but the test between the additive and the saturated model was not significant ($\chi^2(4) = 8.5, p = 0.075$). Thus, for the blue and green conditions, the additive model describes the data most parsimoniously.

Examining the red and yellow experiment in more detail indicated that the significance of the interaction term could be traced to the chroma judgments for the red hue condition. Though the coefficient for the interaction term was small for this condition (0.034 when the slope of the chromatic component was 1.1), it was highly significant ($z = 3.36, p = 0.0008$). This compares with the small and not significant interaction coefficients in the other three conditions for this experiment (chroma task, yellow hue: 0.004, $p = 0.66$; lightness task, chroma red: -0.010, $p = 0.40$; lightness task, chroma yellow: 0.004, $p = 0.76$). Thus, the additive model suffices to describe the data in these three conditions. The lines drawn through the data points in Figure 3.3 are those based on the estimated slopes from the additive model of the GLMM analysis and appear to describe the average data very well. That the line describes the average data well for the red chroma condition also supports that the significant interaction observed in this condition is not very large.

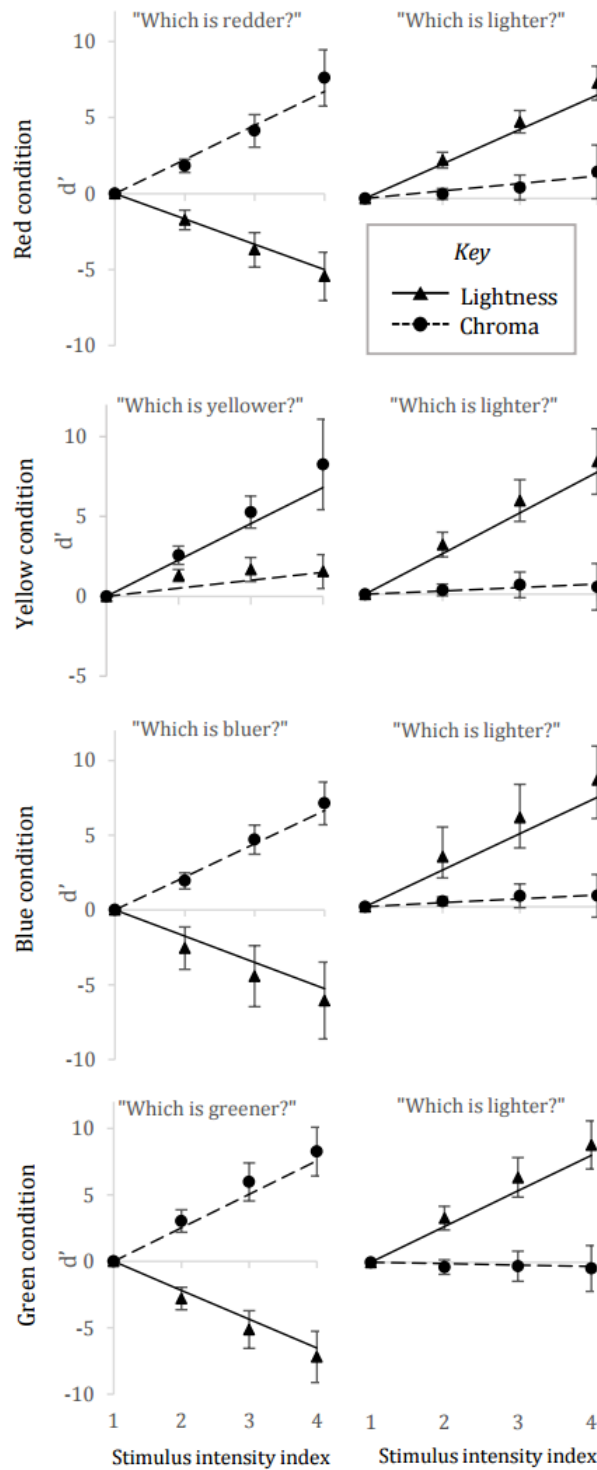


Figure 3.3. These graphs show the normalized parameter estimates for the additive model, averaged across all observers in judgments. Rows show type of judgment, i.e. lightness judgment ("which is lighter?") or chroma judgment ("which is redder/yellower/bluer/greener?"). Columns show the different hue conditions, i.e., red,

yellow, green and blue. Parameter estimates of the contribution of chroma to judgments are indicated by circles and contributions of lightness to judgments is indicated by triangles. The lines show GLMM fit to the data. Error bars show ± 1 SEM.

3.9 Discussion

In this study, we used Maximum Likelihood Conjoint Measurement and Maximum Likelihood Difference Scaling to estimate perceptual scales of lightness and chroma for hues red, yellow, green and blue. In experiment 1, we employed MLDS to estimate difference scales for lightness and chroma. We then used these scales to interpolate four perceptually equally spaced points spanning an equal range along all of the dimensions tested, in order to create the stimuli for the following experiment. The fact that the estimated scales vary nonlinearly with separation in LCh_{uv} simply reflects that this space only approximates uniformity for an average observer (Wyszecki & Stiles, 2000). Prior work has questioned the perceptual spacing of uniform CIE colour spaces (e.g., Hunt & Pointer, 2011; Wyszecki & Stiles, 2000), in particular, that CIE LCh_{uv} perceptual differences are only valid over small differences, and are not reliable over larger distances. The differences in CIE LCh_{uv} units between our stimulus intensity levels are all quite large with respect to discrimination experiments. Deviation from linearity in our data is particularly evident at low chroma values for the red hue: observers in our experiment were more sensitive to differences in chroma here than CIE LCh_{uv} space would predict. This may be an example of a crispening effect (Whittle, 1992) in which observers show heightened sensitivity to small stimulus differences near the background level. The effect may be linked to changes in stimulus appearance at those intensity levels. At the lowest chroma level, the stimulus appears grey, whereas by the 2nd and 3rd level the colour is better described as “red” or “pink”. This categorical jump between an achromatic stimulus and a chromatic stimulus may be related to the high difference sensitivity in this area, although differences in sensitivity could also be the source of the categorical jump rather than a result of it.

In experiment 2, we used the technique of MLCM to analyse paired-comparison judgments of lightness and chroma. In separate experiments, we asked observers to judge either the lightness or the chroma of the stimuli, to determine the “contamination” of one dimension when making judgments about the other. Observers were split for data collection efficiency and the groups were similar in age and gender makeup. We also screened for colour vision deficiencies. Therefore there was no difference between the groups in any attribute we would expect to impact the findings. The results indicate that the lightness and chroma contributions vary linearly as a function of the stimulus index. This certainly reflects, and is an advantage of, our having pre-scaled the stimuli using the MLDS procedure in order that the stimuli be perceptually equally spaced. We conducted MLDS on judgements of achromatic stimuli to produce the lightness scaling for all hues in the second experiment. The linearity of responses in the MLCM analysis support the idea that this approach was valid. In addition, the linearity of response facilitated the application of a more sophisticated statistical analysis of the contributions of the different dimensions and individual sources of variability.

Observers tended to perceive physically lighter stimuli as less chromatic in the red, blue and green conditions but not in the yellow. One potential explanation for this effect is that yellowness and lightness are both defined as an additive combination of the Long and Medium wavelength sensitive cones, and this may cause the two signals to be more easily confounded or it may be that one masks the other. An alternative explanation is that it is due to the wording of the task. Observers were asked to decide which of two stimuli was “redder/yellower/bluer/greener” (depending on the hue condition) when making a chromatic judgment. This may have added a linguistic effect to the decision, such that the observer’s judgment was influenced by the typical lightness of the focal colour of each term. Focal red, blue and green are medium lightness whereas focal

yellow is a lighter colour (Regier, Kay, & Cook, 2005). This may explain why observers judged lighter stimuli as “yellower” for the yellow hue condition, but darker colours as “redder”, “bluer” and “greener”. One way to investigate this would be to use an alternative instruction that does not refer to colour terms such as “pick the more colourful one”. Asking observers outright to choose the more saturated stimulus is problematic as those inexperienced with colour theory find the concept unintuitive (Burns & Shepp, 1988). Pilot testing revealed that observers were unclear on the definition of “saturation” and often confused this for lightness. It is unclear what measure is accessed through the instructions in this experiment, but it may be captured better by a different colour space. For example, in more modern colour spaces such as CIECAM02 and CAT02 from the CIE (Li, Luo, & Hunt, 2000; Moroney et al., 2002), there are related, but distinct, measures of chromaticness (colourfulness, chroma and saturation).

Another pattern in the data is that the physical chroma of the stimuli tended not to influence observers’ perception of the lightness for yellow, green and blue colours. However, in the case of red, increasing the chroma increased the probability that the observer would choose a stimulus as “lighter”. This may be a manifestation of the Helmholtz-Kohlrausch effect (von Helmholtz & Southall, 1925) in which chroma contributes to perceived brightness. The effect is small here but it is possible that larger effects would be observed with stimuli of higher chroma.

The three nested models applicable to MLCM were evaluated within a generalized linear mixed-effects model. It was found that an additive model best described the data in all conditions except for the condition in which observers judged the chroma of red stimuli. For that condition, the additive model of the physical lightness and chroma

contributions to the chroma judgments did not suffice to describe the data and a model including an interaction term described the pattern of observers' judgments better. The contribution of this term was small, however, and more data over a greater stimulus range should be collected to evaluate if it is playing an important role in perception.

Do the MLCM results address the issue of whether lightness and chroma are separable or integral dimensions (Garner, 1974)? This is a more subtle question. The MLCM method shows that we can isolate two response components associated with manipulation of two (psycho-)physically specified dimensions or scales, that are called lightness and chroma in the CIE LCh_{uv} space. We introduce a psychological dimension to the task by asking the observer to order the stimuli on the strength of their perceived lightness or chroma. When the observer makes a lightness judgment, the response component that depends on the scale lightness increases monotonically with the scale value, suggesting that we are tapping a lightness response, and conversely for a chroma judgment. In some cases, the second component follows monotonically the strength of the other scaled component, as estimated in the additive model. Therefore, we are tempted to attribute these components to the psychological lightness and chroma response contributions to the judgments of the observer. In the case of the lightness judgments, three out of four of the cases studied show the lightness component to be independent of the chroma component. It seems reasonable to conclude that the lightness scale is separable from the chroma scale under these circumstances. In one of the cases for the chroma judgments, simple additivity of the response components does not suffice to adequately explain the judgments. It is reasonable to conclude that under this condition, the two responses are not separable and are, therefore, integral. But what of the cases in which the response components are additive, and how should we

interpret the asymmetry in the decision rule for the combination of components depending on which judgment was made?

On the one hand, the fact that observers cannot judge chroma without the contamination of the secondary scale could be taken as evidence that the two scales are psychologically confounded, and this has been argued to be a signature of integral dimensions (Bimler, 2011; A. M. Brown & Lindsey, 2013; Burns & Shepp, 1988). On the other hand, the fact that the judgments can be shown to depend on two responses that add in a way such that the level of one does not influence the level of the other could be taken as a signature of separability. If the former is assumed, then the asymmetry in performance is taken as evidence for integrality of the dimensions and if the latter, separability.

This study has demonstrated the use of MLCM as a technique for investigating how the psychophysical dimensions of colour interact to influence observers' judgments about colour. The utility of pre-scaling the stimulus dimensions to obtain stimuli of equal perceptual spacing and range was demonstrated in that it yielded linear response functions that were simpler to analyse. The results reveal that lightness and chroma components combine additively or nearly so in chroma judgments, but that lightness judgments are largely independent of chroma level over the stimulus range that we studied. The two exceptions to the above results involved the conditions in which the hue component was red, in which the additive model was rejected for chroma judgments and the chroma component was found to add to the lightness when judging lightness. The latter was taken as evidence for the Helmholtz-Kohrausch phenomenon. Interestingly, increased lightness was found to decrease the appearance of chromatic strength in the stimuli for all of the hues tested except yellow.

This study paves the way for further investigation into the dimensions of colour using MLCM, for example, to further investigate evidence of individual differences in colour perception (e.g. Bimler, Kirkland, & Jameson, 2004), to investigate how infants' preferences for colour dimensions develop (A. M. Brown & Lindsey, 2013), or to study the interaction of colour dimensions with other dimensions, such as texture (Saarela & Landy, 2015).

Chapter 4: Paper 3: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Saliency in Infancy

Rogers, M. R.¹, Knoblauch, K.², & Franklin, A.³

¹The Sussex Colour Group, School of Psychology, University of Sussex, Brighton, UK

²Univ Lyon, Université Claude Bernard Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, 69500 Bron, France

4.1 Abstract

How physical dimensions govern children's perception, language acquisition and cognition is an important question in developmental science. Here we use the psychophysical technique of Maximum Likelihood Conjoint Measurement (MLCM) as a novel approach to investigate how infants combine information distributed along two or more dimensions. MLCM is based on a signal detection model of decision that allows testing of several models of how observers integrate information to make choices. We tested 6-month-old infants' preferential looking to green stimuli that covaried in lightness and chroma and analysed infant preferences using MLCM. The findings show that infant looking is driven primarily by lightness, with darker stimuli having a greater salience than lighter, plus a small but significant positive contribution of chroma. This study demonstrates that the technique of MLCM can be used in conjunction with preferential looking to investigate the salience of physical dimensions during development. The technique could now be applied to investigate the role of physical dimensions in key aspects of perceptual and cognitive development such as face recognition, language acquisition and object recognition.

4.2 Introduction

Our world is multidimensional. Objects vary along multiple physical dimensions, for example, apples vary in their glossiness, colour, shape, size and texture. Dimensional properties of an object covary and interact to produce cohesive perceptual experiences. However, it can be difficult to identify how physical dimensions govern perception, cognition and behaviour. This is important to understand perceptual and cognitive development such as language acquisition, object individuation, aesthetics and memory.

Perceptual salience of physical dimensions

During development, some dimensions are particularly salient in a given context. In language acquisition, children exhibit certain attentional biases that lead them to prioritise learning about one dimension over another, depending on the circumstance. For example, in an object label extension task, children display a “shape bias” (Landau, Smith, & Jones, 1988). When children learn a novel word, they are more likely to extend that word to similarly *shaped* objects, versus objects of a similar colour (Bornstein, 1985a), a similar texture or colour (Diesendruck & Bloom, 2003), or a similar size or texture (Landau et al., 1988). However, there is also evidence that such attentional biases are task-dependent, for example, children are more likely to extend a novel word when an object shares a unique *function* than shape (Diesendruck, Markson, & Bloom, 2003). The most salient dimension to a child during novel word learning can be context-dependent. These examples demonstrate the importance of understanding dimensional salience in the domain of language, and having an effective task to test this.

Dimensional properties also contribute to object recognition and individuation during development. In occlusion events, where an object is hidden behind a screen, infants are faced with the challenge of *object individuation* when the object reappears. They must

determine how many objects are moving in and out of view. Objects have dimensional (or featural) properties, which may help solve this puzzle, such as shape, size, texture and colour. If the re-appearing object has different properties to the original one, an adult would be expected to assume that this object is distinct from the occluded object. The dimensions which infants use to solve this puzzle change over the course of development: infants at 4.5 months use shape and size features, at 7.5 months they additionally use pattern and at 11.5 months they can use colour in object individuation (Wilcox, 1999). Covariation of multiple dimensional properties further aids in object individuation. When the colour *or* luminance of an object varies, infants are not able to individuate the object until 11.5 months (Wilcox, 1999), but with covariation, infants aged 7.5 months are able to complete this task (Woods & Wilcox, 2010). This may be because colour and luminance often covary in our natural visual environment and thus are useful cues for identification.

Interaction of dimensions

The evidence discussed above pitted dimensions against each other and examined which dominated as the most important (salient) in a given task. In reality, it is possible that when multiple physical dimensions covary, perception and behaviour arise from an interaction between the dimensions. For example, in adults, the dimensions of colour (hue, chroma and lightness) interact in perception (Burns & Shepp, 1988; Rogers et al., 2016). However, a study which modelled the data from five studies on infants' preferential looking at colours found that 6-month-old infants tended to base their looking preference on hue, and luminance differences did not contribute to the fit of the model (A. M. Brown & Lindsey, 2013). This raises the interesting question of whether there are developmental differences in how the dimensions of colour are weighted in perceptual judgements.

It has also been argued that infants process stimuli more holistically than adults do, and that there is a developmental trend from interaction towards separation of dimensions (e.g. Kemler, 1983; Kemler & Smith, 1978, 1979, Smith, 1979, 1983; Smith & Kemler, 1978). That is, dimensions which adults separate (such as shape and colour), are more integrated in infants' perception. However, research into this hypothesis has yielded mixed results. For example, one study found that younger children (aged 5 to 6 years) did not use holistic rules on a dimensional card-sorting task to a greater extent than older children (aged 10 to 11 years; Kemler & Smith, 1978).

The most appropriate and effective method to investigate dimensional processing in infants is not clear. A method is needed which can quantify the relative contributions of multiple dimensions to perceptual behaviours (e.g., preferential looking) in infancy. Studies conducted with adults on the interaction of dimensions require explicit judgements, such as odd-one-out tasks or similarity ratings (e.g. Burns & Shepp, 1988; Indow & Kanazawa, 1960; Komarova & Jameson, 2013), but these explicit judgement methods are not feasible for infants. In infants, modelling preferential looking with regressions is useful in identifying which dimension contributes to preferential looking (e.g., Brown and Lindsey, 2013), but the method does not address the extent to which multi-dimensional interactions may occur.

Furthermore, when determining how two (or more) physical dimensions contribute to behaviour, there is a challenge of equating the dimensions on a perceptual scale. For example, at 9 months infants appear to identify objects based on shape but not colour (e.g., Káldy & Leslie, 2003), but how can we be sure that the shapes selected in the experiment were equally salient to the colours? This has been addressed using Interdimensional Salience Mapping (ISM; Kaldy & Blaser, 2009) which can determine

how physical change along a dimension relates to changes in saliency, and thus enables saliency to be equated across dimensions. ISM uses a forced-choice preferential-looking method to determine which of two stimuli are more salient in a head-to-head competition. By manipulating the properties of the stimuli and determining which of the two competitors “wins” in such a task, ISM can produce a psychometric function of saliency. This precise approach to stimulus calibration across dimensions is a big improvement on the majority of previous developmental studies of dimensional processing, which do not consider stimulus calibration issues. However, although ISM can equate perceptual saliency across dimensions, we still lack a method that quantifies dimensional *interactions* in perception.

Maximum Likelihood Conjoint Measurement

Here we present a novel approach to investigate the interaction of colour dimensions in infancy. Maximum Likelihood Conjoint Measurement (MLCM) is a psychophysical technique that can quantify the contribution of more than one dimension to a behaviour (Ho, Landy, & Maloney, 2008; Knoblauch & Maloney, 2012). MLCM determines how the probability of a choice between a pair of stimuli is influenced by the covariation along multiple dimensions or attributes of the stimuli. Previous adult studies using MLCM have examined the interactions of: gloss and surface texture (Ho et al., 2008; Qi et al., 2015); surface lightness and gloss (Hansmann-Roth & Mamassian, 2017); contour curvature and luminance in the illusory watercolor effect (Gerardin et al., 2014); the voice and the face in gender perception (Abbatecola, Gerardin, Knoblauch, & Kennedy, 2016); race on the perceived lightness of faces (Nichiporuk, Knoblauch, Abbatecola, & Shevell, 2017); and lightness and chroma in adults (Rogers et al., 2016). In each of these experiments, adult participants made explicit judgements about the stimuli.

However, MLCM has not yet been used to measure the interaction between dimensions

in infancy and throughout development. To implement this approach in infants, we will use preferential eye-movements on presentation of a pair of stimuli as the choice response required to implement MLCM.

In MLCM, the choice probabilities are modelled by a noise contaminated decision rule. This constitutes a signal detection model whose parameters are estimated via maximum likelihood, hence the name of the method. Three nested decision rules can be defined. First, the *independent model* describes the case where the choice probabilities can be described by physical manipulation of one dimension alone; there is no ‘contamination’ by the other dimension. The *additive model* describes the case where choice probabilities depend on the additive sum of two underlying, response components, one associated with each of the manipulated dimensions. Finally, in the *saturated model*, the choice probabilities depend on an interaction beyond the additive combination of the underlying components. Thus, like an interaction in an analysis of variance, the probabilities depend on the particular pair of values along each dimension rather than the sum of two components. The model is called saturated because it includes the maximum number of parameters to model the data. In each case, the decision variable is assumed to be perturbed on a trial-by-trial basis by mean zero, Gaussian noise. This allows for a stochastic relation between the responses and the decision variable. The noise contaminated decision variable is related to the response through a cumulative Gaussian psychometric function.

Current study

In the current study, we use MLCM to estimate the influence of lightness and chroma on infants’ preferential looking to colour. The study aims to establish whether it is possible to use MLCM to model the contribution of physical dimensions to infants’

preferential looking using the technique. We use colour as a testing ground for the application of MLCM to infant preferential looking, and to further our understanding of infants' colour perception. By at least 3 months of age, infants are trichromatic (the three types of cone photoreceptors and the 'red-green' and 'blue-yellow' neural pathways are functioning; Banks & Bennett, 1988; Knoblauch, Bieber, & Werner, 1998; Morrone, Burr, & Fiorentini, 1990, 1993; Volbrecht & Werner, 1987). The ability to discriminate colours progressively improves through development until adolescence (Knoblauch et al., 2001). Previous studies have shown that infants look longer at some colours than others (e.g. Franklin et al., 2008; Franklin, Bevis, Ling, & Hurlbert, 2010; Zemach, Chang, & Teller, 2007) and can categorise the spectrum of colour (e.g., Skelton, Catchpole, Abbott, & Franklin, 2017). However, we know little about how the dimensions of colour contribute and interact in infants' perception of colour. The appearance of colour can be described in a three dimensional perceptual space consisting of *hue*, *lightness* and *chroma* (Wyszecki & Stiles, 2000). As discussed, these dimensions are not independent in adults' perception (Burns & Shepp, 1988; Rogers et al., 2016). It is unknown whether this is the case for infants as well, or whether there are developmental differences in the interaction of colour dimensions.

To test whether MLCM can be applied to infants' preferential looking and to further understand infant colour perception, we conducted an experiment using a forced-choice preferential looking method with 6-month-olds and adult observers (Teller, 1979). The method involves eye-tracking observers' responses to pairs of stimuli, and coding which of the pair they look at first. For infants, the pairs of stimuli were randomly selected from a 3-by-3 matrix of green stimuli in which lightness and chroma independently covary. All stimulus levels were above threshold for infants at 6 months according to threshold data from Knoblauch et al. (2001). The adult experiment used the same range

of lightness and chroma, but the range was divided into four stimulus levels rather than three. We were able to use more stimulus levels with adults as they have better colour discrimination and can tolerate a longer experimental procedure.

We analysed the data using MLCM to estimate the relative contributions of lightness and chroma to the observers' decisions (first look). The outcome variable of our MLCM analysis is parameterised to be on the scale of d' (units of the standard deviation for each scale value) from Signal Detection Theory (e.g. Gerardin, Devnick, Dojat, & Knoblauch, 2014; Green & Swets, 1966; Ho, Landy, & Maloney, 2008; Rogers, Knoblauch, & Franklin, 2016; Stanislaw & Todorov, 1999). Here, we take d' to reflect the perceptual saliency of a stimulus, by measuring which of two stimuli observers look at first. We follow Kaldy and Blaser's definition of saliency as, "the visual system's real-time assessment of the behavioural relevance (current importance) of information in the scene—a prioritization that drives attention allocation and consequent eye movements" (2009, p. 223). The most salient object in a scene is the one that is preferred, i.e. it beats the other in a forced-choice looking paradigm. In our analysis, negative values of d' are possible, as they represent a negative contribution of the relevant attribute to saliency, where saliency is defined by the frequency of first-looks. With MLCM, we can quantify multi-dimensional contributions to perceptual saliency.

4.3 Method

Participants

Twenty-two 6-month olds (9 females) participated in this study in total. One further infant was recruited but did not take part due to fussiness. The infants had a mean age of 28 weeks ($SD = 1.9$) and all had a birth weight greater than 2.5kg, with no family history of colour deficiency. Infant participants were recruited by contacting parents/carers with infant children through the Sussex Baby Lab (University of Sussex, UK). They received a small gift (book or T-shirt) as a thank you at the end of the experiment.

Additionally, twelve adult observers participated (all female) with a mean age of 22 years ($SD = 3.5$). All adult observers were assessed as having normal colour vision using Ishihara plates (Ishihara, 2010) and the Lanthony Tritan Album (Lanthony, 1998). Adult participants were paid £8 per hour for their participation. Ethical approval was obtained from the Sciences and Technology Cross-Schools Ethical Committee at University of Sussex, and the European Research Council Executive Agency ethics committee.

Stimuli and apparatus

Three levels of lightness and three levels of chroma were selected for the infant stimuli, giving a 3-by-3 stimulus matrix (see Table 4.1. Lightness and chroma values for adult and infant stimuli in CIE LCH_{uv} . Hue angle = 143.2° . Table 4.1 and Figure 4.1). The levels were specified in CIE LCH_{uv} colour space. This is a transformed version of CIELUV space, where L is the Lightness, C is the chroma, and H is the hue (Poynton, 2012). This colour space was selected to be in line with a previous study of MLCM in

adults (Rogers et al., 2016). The hue angle was fixed at 143.2° (CIE H), which normal adult observers classify as “green”.

We maximized the stimulus range within monitor gamut (i.e. the possible range of colours than can be displayed on the screen) in order to maximize discriminability for infants. Equal perceptual spacing in CIE LCH_{uv} space between adjacent stimuli was calculated from previously obtained results from adults by Maximum Likelihood Difference Scaling (MLDS) (Rogers et al., 2016). MLDS is a psychophysical scaling technique, also based on a signal detection model of decision, that allows the estimation of an interval scale along a continuous dimension through comparisons of stimulus intervals (Knoblauch & Maloney, 2008; Maloney & Yang, 2003). The derived scale has the property that equal scale differences appear equally different perceptually. The adult stimuli used the same range of values, but divided the range into four levels, giving a 4-by-4 stimulus matrix. A grey background (xyY (1931): 0.313 0.329, 50; $L^* = 100$) was used throughout the experiment, lighter than all stimulus levels.

Table 4.1. Lightness and chroma values for adult and infant stimuli in CIE LCH_{uv}. Hue angle = 143.2°. For the stimuli, these levels of chroma and lightness are combined in all pairwise combinations, form a 3 by 3 matrix for the infant experiment and a 4 by 4 matrix for the adult experiment.

Infant stimuli			Adult stimuli		
<u>Stimulus</u>	<u>Chroma</u>	<u>Lightness</u>	<u>Stimulus</u>	<u>Chroma</u>	<u>Lightness</u>
<u>level</u>			<u>level</u>		
1	5.00	39.63	1	5.00	39.63
2	26.83	52.55	2	19.11	49.55
3	50.00	69.06	3	34.85	58.47
			4	50.00	69.06

Stimuli were presented on a 22-inch Mitsubishi Diamond Plus 230SB monitor, calibrated using a ColorCAL colorimeter (Cambridge Research Systems). Eye-movements were recorded with an EyeLink 1000 Plus Eye Tracker manufactured by SR Research Ltd, using an infant lens. Participants wore a small target sticker, which aids accurate tracking with a freestanding eye-tracker. The experimental procedure was created using SR Research Experiment Builder.

Design and procedure

On each trial, a pair of stimuli was chosen from the 3-by-3 stimulus matrix (see Figure 4.1). All possible unordered pairs were shown in a randomised order, excluding self-comparisons. From the nine stimuli, there are 72 pair combinations, including both left/right and right/left versions of a pair ($9 \times (9 - 1) = 72$).

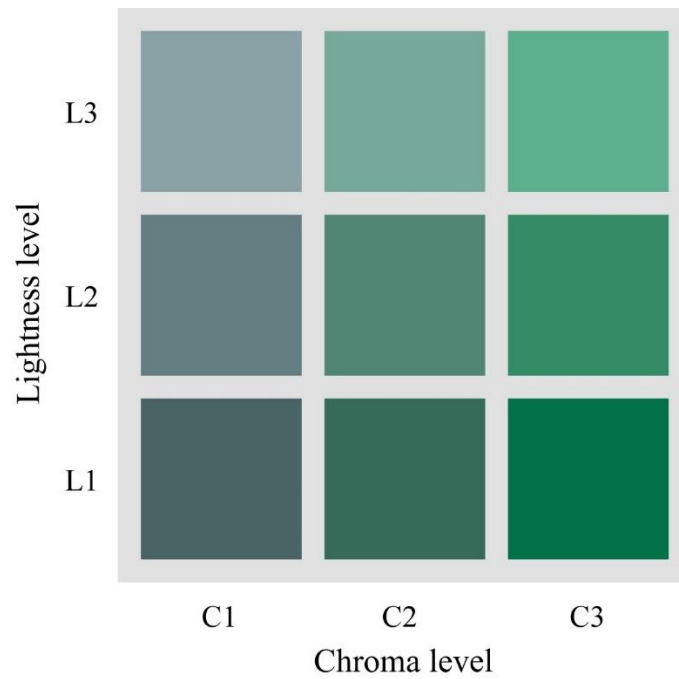
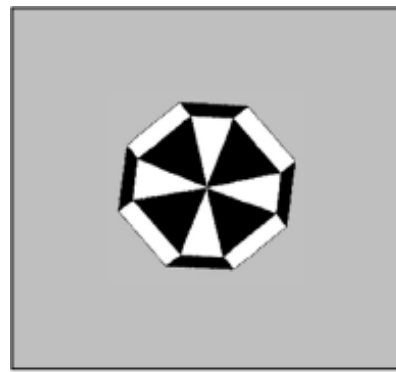
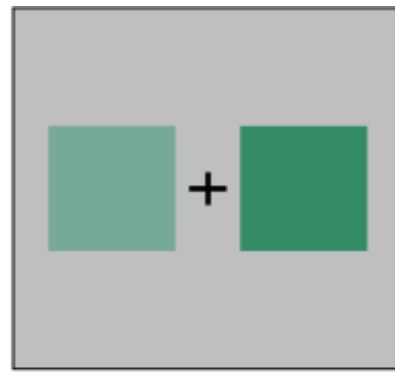


Figure 4.1. The 3-by-3 matrix of infant stimuli. Three levels of chroma and three levels of lightness are perceptually equally spaced using MLDS.

Infants sat in a car seat 60cm away from the computer monitor. Cartoon clips played onscreen whilst the researcher set-up the eye-tracking camera; infants then completed a 3-point calibration. On each trial, a pair of stimuli were presented: one on the left and one on the right hand side of the screen. Each stimulus was 14cm x 14cm on the monitor, corresponding to a visual angle of 13° . If the infant did not fixate on either the left or the right-hand stimulus before the trial timed out (2000ms), the stimulus pair was presented again later in the experiment. Thus, there were, in principle, 72 unique pairs of colours presented, but the number of total trials for an individual observer could be higher as a pair of stimuli would be re-shown until the infant fixated on one side.



Gaze-contingent attention-grabber



Stimulus pair. Experiment proceeds after fixation on one stimulus. Time out after 2000ms.

Figure 4.2. Trial procedure.

To begin, a looming black and white “attention-grabber” appeared with a noise, to ensure the infants were centrally fixated before the trial began. The trial began automatically when the participant fixated on the attention-grabber. The experiment automatically proceeded to the next screen when the infant fixated on either the left or the right stimulus for 160ms (this fixation duration has been found to be reliable and resistant to measurement artefacts; Wass, Smith, & Johnson, 2013). See Figure 4.2 for illustration of trial procedure. The experiment was halted if the infant showed signs of distress. Infants completed an average of 68% of the total number of trials ($SD = 29\%$).

Adult participants were informed that they were a comparison group for an infant experiment, and to “look at the patch that stands out more, or that most grabs your attention”. They completed a 9-point calibration before the trial began. The procedure was the same as the infant experiment, except there was a central fixation cross instead of the attention-grabber. Furthermore, as adults viewed unordered pairs randomly selected from a 4-by-4 stimulus matrix (i.e. 16 stimuli), they responded to 240 trials ($16 \times (16 - 1) = 240$).

4.4 Data analysis

For each participant and on each trial, the stimulus looked at first was recorded as the choice. The data were analysed using Maximum Likelihood Conjoint Measurement; this analysis is described in full elsewhere (Knoblauch & Maloney, 2012; Rogers et al., 2016), and we provide a brief summary here. On each trial, a pair of stimuli are randomly selected from the stimulus matrix (Figure 4.1) for presentation. The pair can be indexed in terms of their ordinal chroma levels (w, y) and lightness levels (x, z). For example, in the pair C3L1 (bottom right stimulus) plus C2L2 (centre stimulus), $w = 3, y = 2, x = 1$ and $z = 2$. It is assumed that when viewing the stimulus pair, the observer forms the noise-contaminated decision variable:

$$\Delta = \delta + \varepsilon = \Psi(\varphi_y^C, \varphi_z^L) - \Psi(\varphi_w^C, \varphi_x^L) + \varepsilon,$$

Equation 4.1

where $\delta = \Psi(\varphi_y^C, \varphi_z^L) - \Psi(\varphi_w^C, \varphi_x^L)$ and the Ψ terms are internal responses that depend on the contributions of the stimulus lightness and chroma to the perceptual saliency. The value of Δ predicts the first-look response, i.e. the observer looks left if it is greater than 0 and right if it is less. The noise term, ε , assumed to be zero-mean Gaussian with variance σ^2 , is included in order to account for the fact that observers do not necessarily make the same choice on repeated trials. MLCM is an equal variance, Gaussian, signal detection model and allows estimation of the scale values corresponding to the contributions of each internal response by maximum likelihood so that they best predict the observer's behaviour over the set of experimental responses.

As described above, there are three possible nested models of the decision variable that can be fit to the data: *independent*, *additive* and *saturated*. With the independent model,

the observer's responses depend on only one of the dimensions. The decision rule reduces to

$$\Delta(w, x, y, z) = \Psi_y^C - \Psi_w^C + \varepsilon$$

Equation 4.2

in the case of a chroma response or

$$\Delta(w, x, y, z) = \Psi_z^L - \Psi_x^L + \varepsilon$$

Equation 4.3

in the case of lightness. With k levels of each dimension, there are $k + 1$ parameters including the variance for the noise term. However, in order to make the model identifiable, the lowest level is set to 0 and the variance to 1, so that there are only $k - 1$ free parameters to estimate.

For the additive model, the decision variable becomes

$$\Delta(w, x, y, z) = (\Psi_y^C + \Psi_z^L) - (\Psi_w^C + \Psi_x^L) + \varepsilon.$$

Equation 4.4

With k levels along each dimension, there are $2k + 1$ parameters, including the variance for the noise term. In order to make the model identifiable, the two lowest levels along each dimension are set to 0 and the variance to 1, yielding $2k - 2$ free parameters to estimate. Finally for the saturated model, the decision variable becomes

$$\Delta(w, x, y, z) = (\Psi_y^C + \Psi_z^L + \Psi_{yz}^{CL}) - (\Psi_w^C + \Psi_x^L + \Psi_{wx}^{CL}) + \varepsilon.$$

Equation 4.5

Due to the interaction terms, the responses cannot be explained by a simple additive sum of component responses, but depend on the specific levels along each dimension. With k levels along each dimension, there are $k^2 + 1$ parameters, including the variance for the noise term. In order to make the model identifiable, only one cell in the $k \times k$ grid is set to zero and the variance to one, yielding a model with $k^2 - 1$ free parameters to estimate. This is the maximum number, hence the term saturated.

If we denote the chroma and lightness quadruple of indices (w, x, y, z) for a trial by q , then assigning responses, R , the values 0/1 to choices left/right, respectively, the probability of choosing the right-hand stimulus on a trial can be written as

$$P(R = 1) = \Phi\left(\frac{\delta_q}{2}\right)^{R_q} \left(1 - \Phi\left(\frac{\delta_q}{2}\right)\right)^{1-R_q}$$

Equation 4.6

where Φ is the cumulative distribution function of the standard Gaussian, the value of 2 in the denominator assures a unit variance for each value of Ψ , and the log likelihood of the set of responses over all trials is

$$\ell(\Psi; R, q) = \sum_q \log(P_q).$$

Equation 4.7

We choose the values of Ψ to maximize the likelihood over the set of responses of the observer by minimizing the negative of the expression above for each of the three models. In practice, the model can be reformulated as a Generalized Linear Model (GLM) with a binomial family (McCullagh & Nelder, 1989),

$$g(E[R = 1]) = X\beta,$$

Equation 4.8

where the link function, g , is the inverse cumulative Gaussian distribution function, X , is a design matrix with a column for each identifiable term in the model filled with 0, -1, or 1 depending on the stimulus levels on the trial and the sign of the coefficients in the decision rule and β is a vector of the estimated identifiable scale values. The three nested models are tested with likelihood ratio tests. The statistic

$$-2(\ell_0 - \ell_1) = \chi^2_{df},$$

Equation 4.9

where ℓ_0 and ℓ_1 are maximum likelihoods from nested models, is distributed asymptotically according to a χ^2 distribution with degrees of freedom the difference of number of parameters between the two models (Wood, 2015). We fit the models to the data and tested them using the MLCM package in the open-source software R (Knoblauch & Maloney, 2014; R Core Team, 2017) that implements the above GLM and testing procedures.

4.5 Results

Infants

Six of the 22 infants completed all 72 trials, and eleven completed over 75% of the trials (range of trials completed = 13 – 72). The analysis of one infant observer (number 8) showed complete separation. This occurs where the outcome variable is perfectly predicted by an explanatory variable (Albert & Anderson, 1984). Further investigation showed that in trials where the highest chromatic value was presented, the infant fixated on this stimulus 100% of the time. While this appears to be prima facie evidence that the infant responded to chromatic stimuli, it presents a problem for analysis because we cannot obtain a valid estimate of the dimension's contribution without variation in response. Therefore, we (conservatively) excluded this infant's data from the full analysis. Analyses reported here are based on data from 21 infants.

The infants' frequencies of first-look responses to the green stimuli are plotted in Figure 4.3. The figure shows that within each chroma level, darker stimuli tend to be more preferred on the first look than lighter stimuli. There also appears to be an effect of chroma, with higher chroma levels slightly more preferred, especially at the highest and lowest lightness levels. A log-linear analysis of the frequencies shows the chroma:lightness interaction fails to reach significance ($\chi^2(4) = 0.86, p = 0.93$), nor does the main effect of chroma ($\chi^2(2) = 0.09$). Nevertheless, the model that included main effects of lightness and chroma displayed a lower AIC than one without the chroma term ($AIC(L + C) = 70.8$; $AIC(L) = 71.5$) suggesting that the model with the chroma term is closer to the true model.

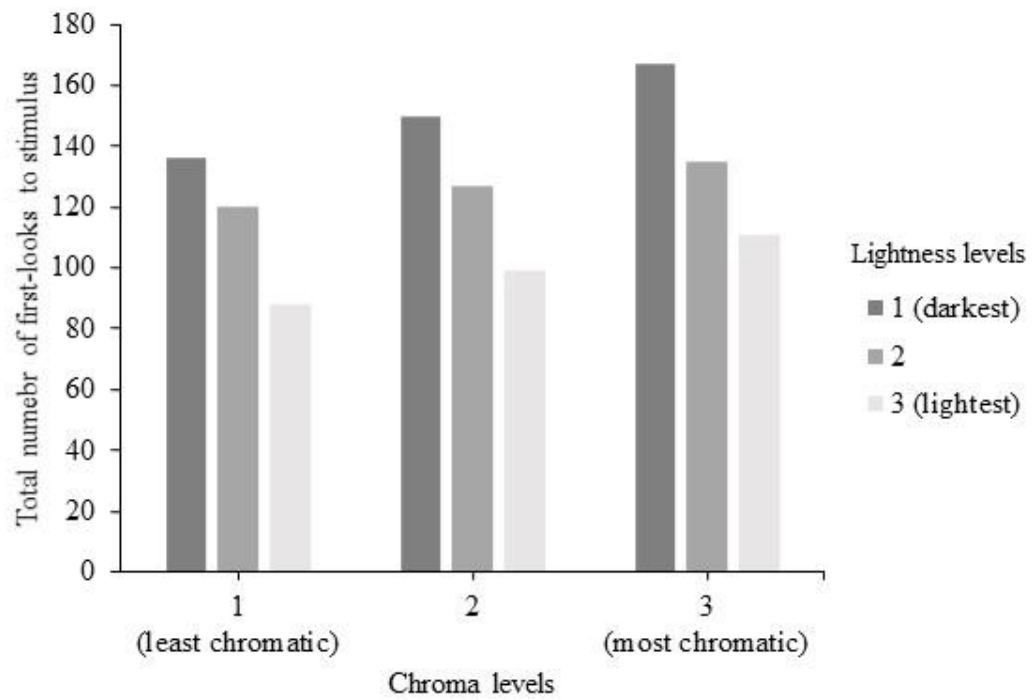


Figure 4.3. Frequency of first-look responses to each of the stimuli in the 3-by-3 stimulus matrix, summed across all infant observers. $N = 21$.

The above results were obtained by combining responses across all infants. We used MLCM to fit the additive model to the responses of each individual infant to estimate the contributions of each dimension to the looking preferences. The points in Figure 4.4 show the average additive contributions of lightness and chroma to infants' looking preference and the 95% confidence intervals. The y-axis indicates d' , which is a measure of sensitivity used in psychophysics, and that here indicates perceptual salience. Negative values of d' indicate a negative contribution of the relevant attribute to the frequency of first-looks. Figure 4.4 shows that lightness negatively contributes to infants' first-look preference, i.e., a darker stimulus tends to be preferred over a lighter stimulus. Conversely, chroma shows a smaller positive contribution to looking preference, indicating that stimuli with higher chromaticity are preferred over stimuli with lower chromaticity. The responses increase approximately linearly with the stimulus indices, suggesting that the stimulus spacing based on the adult scaling functions holds approximately for the infants, as well. While the estimated values of d' are small, the confidence intervals for the strongest lightness and chroma values indicate that they do differ significantly from zero.

A criticism that can be raised is that the number of trials to estimate the response probabilities is quite small. If all 72 trials were completed, then there are only 2 trials for each of the 36 unique pairs. As indicated earlier, not all infants completed 72 trials and one infant completed as few as 13, surely insufficient to estimate all of the choice probabilities. For comparison, previous studies in adults that used more levels for each dimension tested on the order of 1000 or more trials. Knoblauch & Maloney (2012) have shown with simulation that the precision of estimates in MLCM is related to the square root of the number of trials tested. Comparing, for example, with the Gerardin et al. (2014) study that tested 1500 trials per condition, we would see a reduction of

precision of $\sqrt{72/1500} = 0.22$, at best, compared to their results. That is, we expect the estimation of the scale values for 72 trials compared to 1500 trials, to be about 4-5 times worse (error bars 4-5 times larger).

Mixed-effects models provide a possible solution to this problem in that observers are assumed to be sampled from a common population that shares common characteristics (Knoblauch & Maloney, 2012; Moscatelli et al., 2012). Information is pooled over all observers, weighted according to the information available from each observer's data, to obtain an optimal population estimate. This leads to shrinkage of the predictions of extreme observer's means towards the population mean, as the individual estimates are considered to borrow strength from each other.

Generalized Linear Mixed-effects Models (GLMM) are GLMs in which the linear predictors are composed of fixed and random effect terms. Estimates are made of the fixed-effect coefficients and the variances of the random terms (Bates et al., 2014). We can use the approximate linearity of the response estimates to specify a GLMM for MLCM (Rogers et al., 2016) that combines the data from all infants in an optimal fashion. We are assuming that this equal spacing based on difference scaling is also applicable to infants, and therefore the assumption of linearity in response is valid. In short, for each stimulus pair, we compute the difference of levels within each dimension, dC and dL, and use these variables as covariates. The GLMM can then be expressed as

$$g(E(R = 1)) = (\beta_C + b_{C,o})dC + (\beta_L + b_{L,o})dL + \epsilon,$$

Equation 4.10

where g is the probit (inverse cumulative Gaussian) link function, β_C and β_L , fixed effect slopes for chromatic and luminance contributions, $b_{C,o}$ and $b_{L,o}$, observer specific random variations of the slope each assumed to be Gaussian random variables with $\mu = 0$ and variances σ_C^2 and σ_L^2 , respectively, and ϵ is a standard Gaussian variable with $\mu = 0$ and $\sigma^2 = 1$. The mixed-effects models were fit using the *glmer* function in the lme4 package (Bates et al., 2014) within the software R (R Core Team, 2017). We can use the same framework to fit the independence model by dropping one of the terms, covariates, dC or dL, and a saturated model by including terms that are the product of the two covariates. The three models are then evaluated using nested likelihood ratio tests.

Likelihood tests indicated that the additive model fit the data significantly better than both the independence lightness model ($\chi^2(4) = 5.90, p = .015$) and the independence chroma model ($\chi^2(4) = 19.79, p < .001$). However, the saturated model was not an improvement over the additive model ($\chi^2(5) = 0.289, p = .591$). This suggests that both chroma and lightness contribute to infants' looking preference, but there is not an interaction between the two dimensions. The additive model revealed a significant negative effect of lightness ($\beta_L = -0.202, z = -5.76, p < .001$), and a smaller but still significant positive effect of chroma ($\beta_C = 0.096, z = 2.48, p = .013$) on looking preference. The 95% confidence intervals for the slopes, based on profile likelihoods are: chroma: (0.020, 0.179) and lightness: (-0.275, -0.130), both excluding zero. The variance associated with the chroma term was 2.6 times larger than that for the lightness term, or in other terms, it accounted for 72% of the inter-observer variability. The lines drawn through the data points in Figure 4.4 are based on the estimated slopes from the additive model and appear to describe the average data well, although the points for the chromatic averages suggest some bias in the estimates, perhaps arising from the small

number of samples for the individual estimates. This also supports the use of the equal perceptual spacing of stimulus levels based on adult MLDS data. The envelopes about each curve display the 95% confidence limits on the fitted lines. Again, while the predicted effects are small, they do support significant contributions of both dimensions to the infants' judgments with a contribution of lightness to performance roughly twice that of the contribution of chroma.

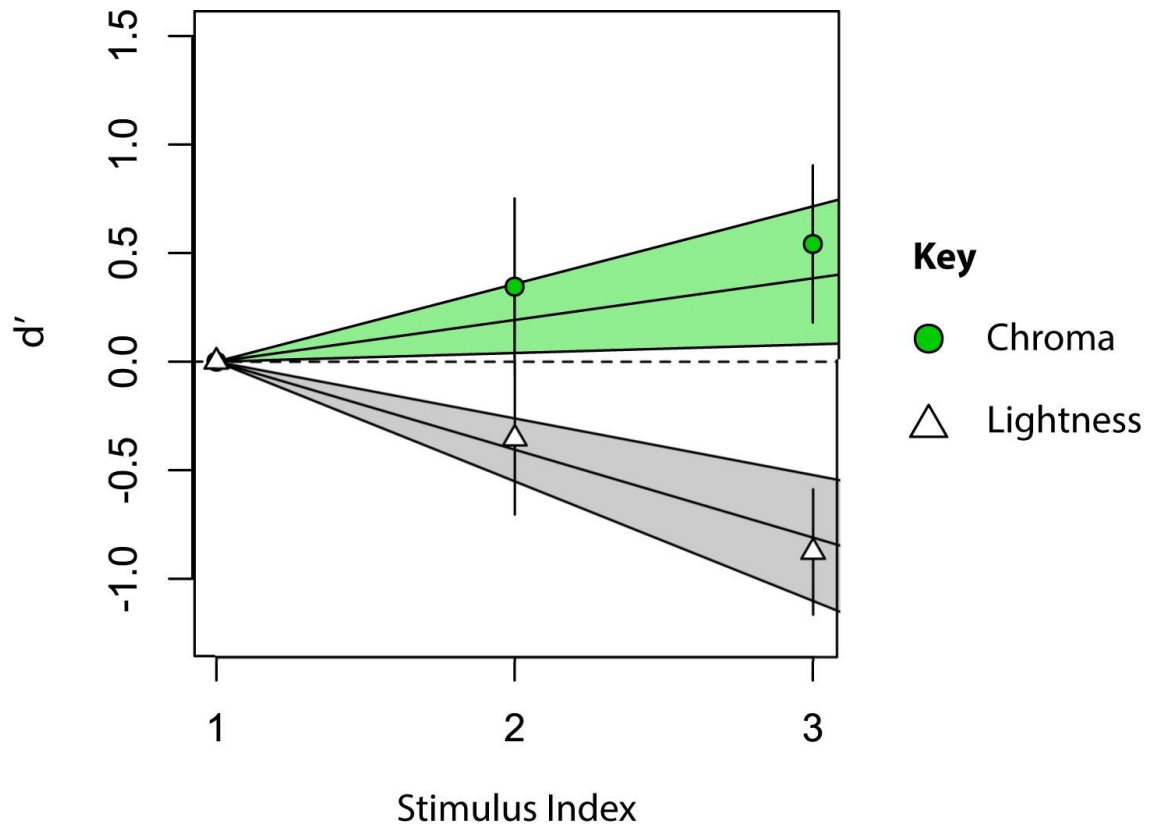


Figure 4.4. Parameter estimates for the additive model for looking preference in infants, averaged across observers. Circles show the estimated contribution of chroma and triangles of lightness based on individual MLCM fits. Error bars show 95% confidence intervals of the points. The lines are based on the mixed-effects model fixed-effect slope estimates, and the shaded areas the 95% confidence intervals for the slopes of the lines. $N = 21$.

Adults

As with the infant observers, we plotted the combined adults' first-look response as a bar chart in Figure 4.5. It appears that adults are more likely to look at a stimulus as it becomes more chromatic. There also seems to be a positive, rather than a negative, influence of lightness, with lighter stimuli being more preferred, especially at the highest lightness level (4) compared to the other three. The log-linear analysis did not support a significant lightness:chroma interaction ($\chi^2(9) = 9.1, p = 0.43$). Both the lightness and chroma main effects were significant, however (lightness: $\chi^2(3) = 59.3, p \ll 0.001$; chroma: $\chi^2(3) = 301.7, p \ll 0.001$).

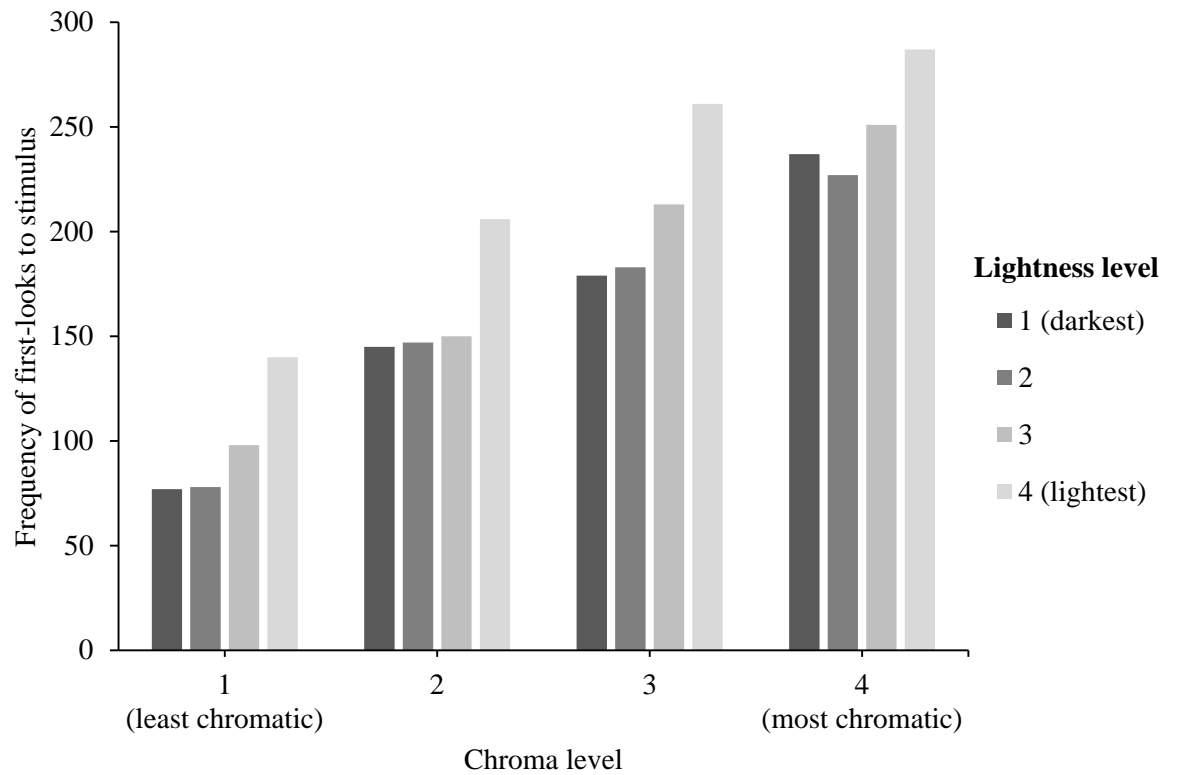


Figure 4.5. Frequency of first-look responses to each of the green stimuli in the 4-by-4 stimulus matrix, summed across adult observers. $N = 12$.

Individual adult responses were analysed using MLCM with the additive model, and the average values for the contributions of lightness and chroma components with 95% confidence intervals are plotted in Figure 4.6. Chroma has a large positive effect on looking preference: adults are more likely to look first towards a stimuli with a high chroma level. Lightness also has a positive contribution to first-look preference in adults, although it appears smaller, and is perhaps significant only at the highest lightness level.

The GLMM described above to test the infant data was fit to the adult responses.

Likelihood tests rejected the independence model for lightness ($\chi^2(4) = 10.94, p < .001$) and for chroma ($\chi^2(4) = 4.63, p = .031$). However, the saturated model fit could not be differentiated statistically from the additive model fit ($\chi^2(5) = 0.017, p = .897$). This suggests that both chroma and lightness contribute additively to adults' looking preference without the need of an interaction term. The mixed-effects analysis revealed a significant positive effect of chroma ($\beta_C = 0.599, z = 4.19, p < .001$), and a smaller but still significant positive effect of lightness ($\beta_L = 0.226, z = 2.37, p = .018$) on looking preference with 95% confidence intervals: chroma: (0.298, 0.908) and lightness: (0.024, 0.432). The lines drawn through the data points in Figure 4.6 are based on the estimated fixed-effects slopes from the additive model of the mixed-effects analysis. The variance associated with the chromatic term was 2.3 times larger than that for the lightness term, or in other terms, it accounted for 70% of the inter-observer variability, similar to the infant value. The magnitude of the average effects are about three times larger than those of the infants. In adults, the contribution of chroma to performance is nearly three times that of lightness, i.e., there is a reversal in the dominant dimension between adult and infant.

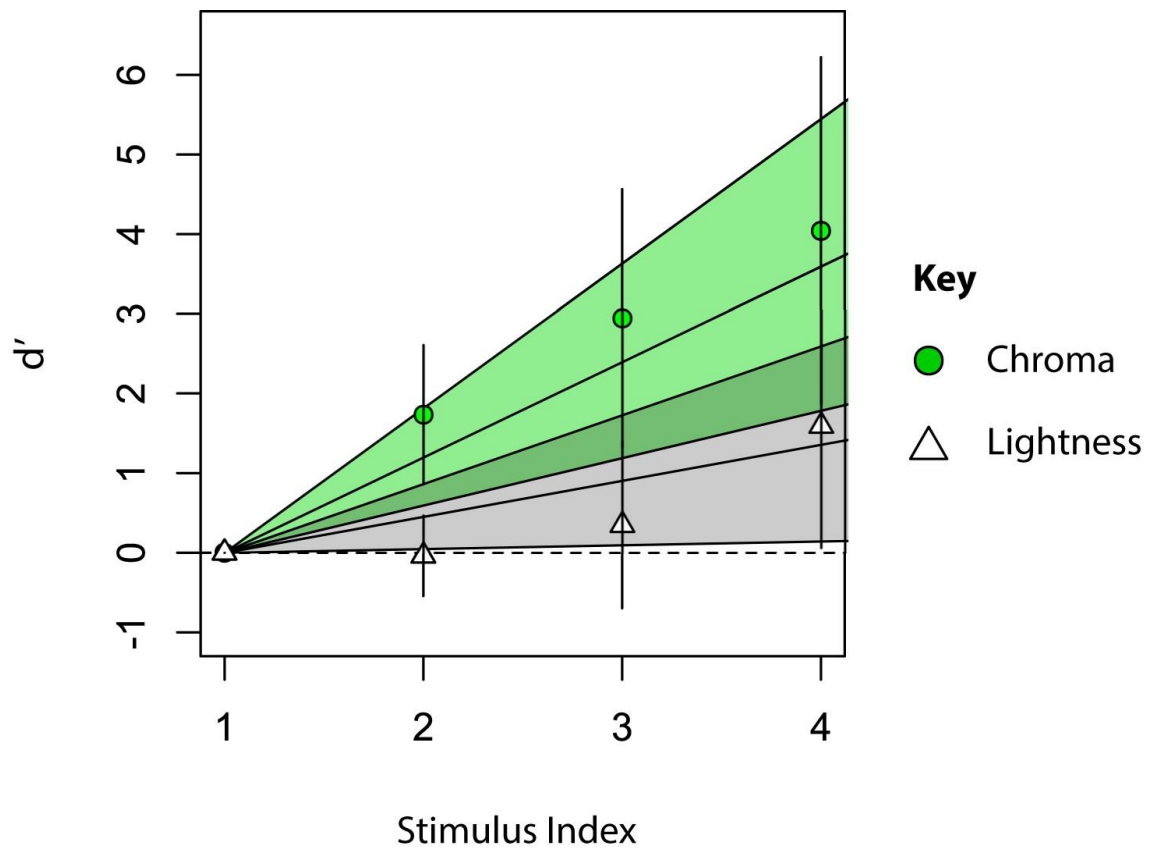


Figure 4.6. Parameter estimates for the additive model for looking preference in adults, averaged across observers. Circles show the estimated contribution of chroma and triangles of lightness. Error bars show 95% confidence intervals. $N = 12$.

4.6 Discussion

This study used the psychophysical technique of Maximum Likelihood Conjoint Measurement to investigate how the dimensions of lightness and chroma contribute to perceptual saliency in infants and adults. Previous studies have demonstrated the effectiveness of this method when adults are instructed to judge pairs on specific stimulus dimensions (Gerardin et al., 2014; Ho et al., 2008; Rogers et al., 2016), but it was still an open question as to whether the methods could be extended to infants. Here, we used eye tracking with infant and adult observers, whilst they viewed pairs of green stimuli that varied in lightness and chroma to obtain first fixations as a choice response measure. We successfully used MLCM to model the first-look data with a generalized linear mixed-effects model.

The analyses reveal that chroma and lightness both contribute to looking preference in infants and adults, and that an additive model best describes the data in both groups. This indicates that the looking response depends on an additive combination of the underlying response components to lightness and chroma, but that there is no interaction between the two components. Infants' looking behaviour was primarily predicted by lightness, but there was a small positive contribution of chroma. Infants prefer stimuli that are darker and more chromatic. Whereas, for adults, chroma primarily determined the first-look response, and there was a small positive contribution of lightness. Adults prefer stimuli that are more chromatic and lighter.

Interestingly, in our former study, when adults were instructed to judge which stimulus is greener, higher chroma led to a more positive chroma contribution but higher lightness to a more negative lightness contribution (Rogers et al., 2016), qualitatively similar to the infants' saliency responses in the current study. This raises the possibility

that the mechanisms engaged by the saliency of the stimulus in the infants are the same as those employed by an adult judging the chromatic difference between a pair of stimuli. However, the results in both studies show that chroma and lightness jointly contribute to these behaviours. In the Rogers et al. (2016) study, when adults were instructed to judge which stimulus was lighter, the lightness component dominated and the response was independent of the chroma response. This is unlike either of the adult or infant response patterns in the current study.

Development of dimensional interaction

Previous researchers have theorised that perceptual dimensions become more separated over development, and that infants process stimuli more holistically than adults do (Kemler, 1983; Kemler & Smith, 1978). However, our findings do not support the idea that there are differences between infants and adults in the extent of dimensional separation for colour, as the additive model best fit both infant and adult data. This suggests that in adults and infants, both lightness and chroma contribute to perceptual saliency additively, and that there is not an interaction between the two dimensions. However, there was a difference between adults and infants in the way lightness and chroma contributed to response. There are a number of possible explanation for these differences.

Visual strategy

First, the visual strategy may be different between adults and infants. We theorise that adults are focusing more on the abstract stimulus property of colour, and ignoring the background, whereas infants are focusing on the whole screen in a trial, including the background. This may lead to luminance contrast driving infant looking behaviour, which would bias the infants towards lightness-based responses. Previous work has

Chapter 4: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy

shown that for achromatic stimuli, 4-month-old infants' looking preferences are governed by luminance contrast (stimulus to background luminance ratio; Chien, Palmer, & Teller, 2005). Here, lightness and chroma levels were perceptually equated using difference scaling data from adults (Rogers et al., 2016), and the fact that the average responses vary approximately linearly as a function of the stimulus index suggests that this scaling was valid. However, all stimuli were darker than the background, and the lightness levels had greater variation from the background than the chroma levels in CIE delta E. Therefore, the pattern of infants' responses may be driven by a preference to stimuli that were most different to the background. If the above 'contrast theory' is correct then infants may fixate more on the edges of the stimuli, whereas adults may focus on the centre. It would be informative to test this theory by displaying the stimulus pairs for a set amount of time and measuring looking time to each area of the screen.

Cognitive strategy

A second possible explanation for the difference in dimensional contribution between infants and adults is a difference in the cognitive strategy between the two groups. Adult participants were given instructions for the task ("look at the patch that stands out more, or that most grabs your attention"). With these instructions, we aimed to access the same outcome we measured in infant participants: salience. However, the instruction is vague. By making the salience outcome measure in adults explicit, we may have inadvertently introduced unwanted cognitive strategies. We may be accessing our target measurement of 'salience' more successfully in infants than in adults. This theory is supported by the large variation in adult responses (see error bars of Figure 4.6), which may indicate different interpretations of the task among adult observers. In comparison, the variation among infants is much smaller. This is despite the fact that adults

performed many more trials than infants (240 for adults compared to a maximum of 72 for infants). Adult eye-movement strategies are highly dependent on the task or instruction given (Buswell, 1935). It would be informative to examine how the dimensions of colour interact to inform looking behaviour in adults with a range of different instructions.

Dimensional differences

A final explanation for the difference in results is that there are differences in the perceptual weighting of the dimensions between infants and adults. In 6-month infants, hue preference curves are highly similar at different levels of lightness (A. M. Brown & Lindsey, 2013). Furthermore, when colour and luminance were equated for visual salience, 6.5-month infants noticed a colour change, but not a luminance change when tracking an occluded object (Kaldy, Blaser, & Leslie, 2006). This may lead to the inference that lightness does not have an effect on infants' looking preference. However, in this study, we find that in fact lightness makes a stronger contribution than chroma to looking preference in infants.

The finding that lightness influences infant preference more than chroma may be due to the relative faster maturation of the magnocellular pathway, compared to the parvocellular pathway, in the first year of life (Hammarrenger et al., 2003). Neurons in the magnocellular pathway are more sensitive to luminance contrast, whereas neurons in the parvocellular pathway are more sensitive to red/green chromatic contrast (B. B. Lee et al., 1990; V. C. Smith et al., 1995). Three- and 4-month infants are more sensitive to luminance contrast than chromatic contrast (D. Allen et al., 1993; Dobkins et al., 1999). There is evidence of adult-like performance of the magnocellular pathway in 4-month infants; whereas the parvocellular pathway had not fully developed by this age

(Dobkins et al., 1999). During infancy, the parvocellular pathway has less cortical input than the magnocellular pathway and has a slower development. Whilst there is some chromatic input to the parvocellular pathway, infants may use this differently from adults due to level of pathway maturation (Dobkins & Anderson, 2002). Our results may be explained by greater sensitivity to lightness differences than chroma differences in our 6-month infant sample, due to greater magnocellular maturity than parvocellular maturity at this age.

Future directions

This study has demonstrated that a simple method in combination with statistical modelling based on a signal detection model allows scaling of the contributions of stimulus dimensions in perceptual saliency in human infants. The method, like most in developmental science, is limited by the patience of the infant participants. In this study, we successfully modelled three levels of lightness and three levels of chroma. Additional levels or dimensions would likely have resulted in too many trials for infants to complete in one session. Importantly, the use of many infant observers in the framework of mixed-models allowed us to overcome the low number of trials recoverable from individual infants. An advantage of the method is that it requires a binary measurement, therefore we were able to exploit first-look responses (left or right stimulus). This means that the trials moved along rapidly. Other measures of looking time, for example, stimulus fixated for the longest duration, could equally be exploited and might yield interesting results.

MLCM modelling assumes that the observer is making a choice, and this is the basis for estimation of the decision variable. For example, in Rogers et al. (2016), adult observers viewed pairs of stimuli that varied in lightness and chroma, and made a judgement

Chapter 4: A Novel Method to Investigate How Dimensions Interact to Inform Perceptual Salience in Infancy

about whether the left or right stimulus was lighter or more chromatic in different sessions. In this study, we use eye-movement data as a proxy measurement for stimulus choice. It is always challenging to determine exactly what is being measured in infant eye-tracking studies (Aslin, 2007). However, eye-movements are widely interpreted as choices in infants across developmental psychology (Civan, Teller, & Palmer, 2005; Teller, 1979).

The results reveal interesting differences in responses to colour between adults and infants, and there is now potential for this method to be applied in a wider range of contexts. This method could be used to study dissociation between visual domain over the course of development (Dobkins, 2009) by pairing dimensions from different modalities such as colour, motion, form, depth. The development of early cross-modal correspondences could also be studied in this way. For example, auditory-visual correspondences such as associations between high-frequency, small size and bright colour (as in Haryu & Kajikawa, 2012).

Conclusion

This study demonstrates how a signal detection method can be used to investigate salience and interaction of dimensions in infancy. The contributions of luminance and chroma to salience judgments in infants resembled the contributions of these dimensions to adults judging stimuli on the basis of chroma differences. This study paves the way for future work that aims to understand the contribution of various perceptual dimensions to perceptual and cognitive development to benefit from this method.

Chapter 5: Paper 4: Why is colour discrimination poorest along the daylight locus?

Rogers, M. R.¹, Lee, R. J.², & Bosten, J. M.³

1 The Sussex Colour Group, School of Psychology, University of Sussex, UK

2 School of Psychology, University of Lincoln, UK

3 School of Psychology, University of Sussex, UK

5.1 Abstract

Previous studies have suggested that illumination discrimination is poorer along the blue-yellow axis compared to other colour directions as an aid to, or because of, colour constancy mechanisms. However, these ideas lack a sufficiently controlled experiment comparing discrimination of the colours of illumination to discrimination of the colours of surfaces. Equally poor discrimination performance in the blue-yellow direction, compared to the red-green direction, for both cues to the illumination and cues to surface colours, would suggest that the blue-yellow bias in colour discrimination performance is not specific to colour constancy. Instead, it may be the result of a range-accuracy trade-off due to the greater range of blue-yellow surfaces and illuminations in natural environments, or some other mechanism. In this study, we rendered stimuli to compare chromatic discrimination thresholds for changes in specularly reflected light (a cue to the illumination) with chromatic discrimination thresholds for changes in diffusely reflected light (a cue to surface reflectance). Our results showed no significant difference between conditions in blue-yellow discrimination relative to red-green discrimination. However, there was a trend towards relatively poorer blue-yellow discrimination for specularly reflected light than for diffusely reflected light.

5.2 Introduction

People with normal trichromatic vision have three different types of cone receptors, sensitive to long (L), medium (M) and short (S) wavelengths of visible light. After light is received by the cones, colour information is thought to be passed through the early visual system via two chromatically opponent mechanisms (Jameson & Hurvich, 1955; Mollon & Cavonius, 1987) which compare the activities of the different cone types. These opponent mechanisms are the basis of the MacLeod-Boynton (1979) chromaticity diagram, a chromaticity space where one opponent mechanisms ($L / (L + M)$) varies along the x-axis, and the other, ($S / (L + M)$) varies along the y-axis.

Based on these early mechanisms of colour vision, we might expect that colour discrimination performance would be aligned with one of the cardinal axes in the MacLeod-Boynton chromaticity diagram. However, previous evidence shows that colour discrimination is actually optimised along an axis that is intermediate to the two cardinal axes. Discrimination in the red-green direction (roughly the positive diagonal in the MacLeod-Boynton chromaticity diagram) is better, relative to discrimination along the blue-yellow direction (Bosten et al., 2015; Hansen, Giesel, & Gegenfurtner, 2008; Krauskopf & Gegenfurtner, 1992; Webster, 1996). Relatively poor discrimination along a blue-yellow axis has also been reported in studies that use colour spaces other than the MacLeod-Boynton diagram, such as CIE u^*v^* (Álvarez, Linhares, Moreira, Lillo, & Nascimento, 2017), CIE Lu^*v^* (Pearce et al., 2014; Radonjic, Pearce, et al., 2016) and CIE xy (Romero, García, del Barco, & Hita, 1993).

Several theories have been proposed to explain this unexpected asymmetric pattern in colour discrimination. One theme unifies these theories: that the anisotropy of colour discrimination performance is related to statistical regularities of colour in the

environment. One such source of statistical regularity in natural environment is illumination, which is most likely to be sunlight, skylight or mixtures thereof (Shepard, 1995). Daylight has a chromaticity that varies more along the blue-yellow direction than along the red-green direction (Foster, Amano, Nascimento, & Foster, 2006; Hernández-Andrés, Romero, Nieves, & Lee, 2001; Philipona & O'Regan, 2006; Webster & Mollon, 1997). This is known as the daylight locus (Wyszecki & Stiles, 2000) of illumination chromaticities.

There also is some limited evidence to suggest that the surface reflectance chromaticities in the natural environment vary more in the blue-yellow direction than in the red-green direction (Bosten et al., 2015; R. O. Brown, 1994; Macleod & von der Twer, 2001). Knowledge about chromatic distributions in natural scenes comes from taking hyperspectral images of reflectance spectra. In such measurements, it is difficult to disentangle surface reflectance from illumination. Therefore, it is difficult to ascertain whether a blue-yellow bias in chromatic distributions originates from variations in daylight solely, or whether there is also a bias in surface reflectances. However, Brown's (1994) hyperspectral measurements taken in San Diego use the true reflectances measured relative to a similarly oriented white diffusing reflectance standard, which normalises out the effect of daylight. In these measurements, variation was lower in the red-green direction than the blue yellow direction (Macleod & von der Twer, 2001).

Colour constancy and the daylight locus

The surface colours of objects in the real world are generally stable. However, illumination chromaticity changes in different environments, in different areas of a scene due to shading and inter-reflections, and over the course of the day. Colour

constancy is the ability of our visual system to base perception on the stable surface reflectance properties of objects, discounting any changes in the illumination spectrum. The light reaching the eye from an object is a product of the illumination spectrum and the object's surface reflectance function (Foster, 2011; Hurlbert, 2007; Smithson, 2005). One way our visual system may keep our perceptual experience of the colour of an object stable and consistent, is to extract the surface reflectance properties from the reflected light reaching the eye, and discount the illumination spectrum. The task of colour constancy is challenging, because the visual system simply receives three cone input signals, and there is no immediately obvious way to disentangle the contributions of light that are a property of the surface reflectance, and light that is a result of the illumination.

Researchers have hypothesised that relatively poorer blue-yellow illumination discrimination is related to colour constancy (Golz & MacLeod, 2002; MacLeod, 2003; Macleod & von der Twer, 2001; Pearce et al., 2014). In what we call in the current paper the *adaptive insensitivity* hypothesis, poor chromatic discrimination along a blue-yellow axis actually improves colour constancy (Pearce et al., 2014; Radonjic, Pearce, et al., 2016). As natural illuminations are aligned along a blue-yellow axis, relative insensitivity to blue or yellow colour changes helps us to ignore changes in natural illuminations.

Alternatively, the *constancy noise* hypothesis (Bosten et al., 2015) proposes that relative insensitivity to colours along a blue-yellow axis arises *because* of colour constancy, rather than as an aid to it. That is, we make internal assumptions about the likelihood of illumination chromaticity (based on our experience with the world) and uncertainty about illumination gives a relatively noisy colour representation along the blue-yellow

axis. Unlike the adaptive insensitivity hypothesis, the constancy noise hypothesis does not predict that constancy would be best along a blue-yellow axis. It proposes that uncertainty about illumination adds noise to colour judgements along that axis, so would increase the variance of achromatic settings along a blue-yellow axis without altering the mean.

Pearce, Radonjic, Hurlbert and colleagues (Pearce et al., 2014; Radonjic, Pearce, et al., 2016), in their *adaptive-insensitivity* hypothesis, have argued that illumination discrimination is a proxy measurement for colour constancy ability, and that poor illumination discrimination indicates good colour constancy. They conducted a forced-choice illumination matching experiment in which participants viewed a target scene lit by an illuminant, followed sequentially by two other scenes, and were asked whether the illumination for scene one or for scene two was more like that of the target (Pearce et al., 2014). Illumination was manipulated in four chromatic directions: red, green, yellow and blue in CIE Lu^*v^* . They found poorest discrimination of bluer illumination changes, and most accurate discrimination for greener illumination changes.

Discrimination of redder and yellower changes were equal. A subsequent study used a similar experimental procedure, but with an adaptive staircase method rather than the method of constant stimuli (Radonjic, Pearce, et al., 2016). They found, again, poorest discrimination for illumination changes in the blue direction, and best discrimination for changes in the green direction.

If poor blue-yellow discrimination were adaptive for colour constancy, we would predict that colour constancy is better along the daylight locus compared to other directions. However, measures of colour constancy (as opposed to discrimination) in this area have produced mixed findings (e.g. Daugirdiene, Kulikowski, Murray, &

Kelly, 2016; Worthey, 1985). Brainard (1998) used an achromatic matching task to test colour constancy in eight hue directions. He found that colour constancy performance did not depend on the chromaticity of the illuminant change: there was no advantage along the daylight locus. However, this study was limited by a sample size of only two observers. Delahunt and Brainard (2004) used an achromatic adjustment task as a measure of colour constancy under illuminations that mimicked natural daylight (blue and yellow) and illuminations which were unnatural (red and green). They found, as expected, that colour constancy was poorest under changes in red illumination. However, colour constancy was best (and comparable) under both green and blue changes in illumination. These mixed findings suggest that the *adaptive-insensitivity* hypothesis may not fully explain non-uniform colour discrimination.

Bosten et al. (2015) based their *constancy-noise* hypothesis on the distribution of settings of unique white (a white that does not contain any red, green, yellow or blue), which varies more along the blue-yellow axis than the red green axis (Beer, Dinca, & MacLeod, 2010; Bosten et al., 2015; Chauhan et al., 2014; Honjyo & Nonaka, 1970; McDermott & Webster, 2012; Webster & Leonard, 2008; Werner & Scheffrin, 1993; Werner & Walraven, 1982). When an observer is required to make a setting of white in complete darkness the stimulus is ambiguous. A yellow stimulus, for example, could be interpreted as a yellow surface under white illumination or as a white surface under yellow illumination. They proposed that uncertainty about the precise colour of the illumination, combined with the assumption that illumination tends to be distributed along the daylight locus, results in the observed variability in white settings. Bosten et al. conducted a series of experiments to test the *constancy-noise* hypothesis.

In these experiments, task type and stimulus surround were manipulated to either give a cue to the illumination or give no cue. If the *constancy-noise* hypothesis is correct, providing a cue to the illumination should reduce variability along the blue-yellow axis. Participants made absolute judgements of white (no comparison stimulus, black background) and relative judgements of white (bipartite stimuli and 4AFC, black background); this should change the influence of uncertainty about the illumination. In a further experiment, participants made relative white judgements when the stimulus surround was either black (no cue to the illumination) or grey (cue to the illumination). The findings were mixed: task type did not affect blue-yellow variability, but a grey background did significantly improve discrimination in the blue-yellow direction compared to a black background. To explain these findings, Bosten et al. proposed an alternative explanation for non-uniform colour discrimination, termed the *range-accuracy trade-off* hypothesis.

The range-accuracy trade-off hypothesis

In the efficient coding hypothesis, visual encoding is optimised to represent the visual information contained in scenes (Laughlin, 1981, 1983). A putative blue-yellow colour mechanism must represent a higher amount of information compared to an orthogonal red-green mechanism, due to the relatively greater blue-yellow chromatic variability contained in natural environments. Therefore, optimal coding of the blue-yellow pathway confers coarser representation, compared to optimal coding along other colour directions. Larger steps in representation result in lower chromatic resolution and reduced chromatic discrimination. This theory of non-uniform colour discrimination is termed the range-accuracy trade-off hypothesis, because the greater distribution of blue-yellow of natural chromaticities in colour-opponent space necessitates that accuracy be

sacrificed for range in the system (Laparra, Jiménez, Camps-Valls, & Malo, 2012; MacLeod, 2003; Macleod & von der Twer, 2001).

To distinguish between the three theories outlined here, an experiment is needed that precisely measures surface discrimination compared to illumination discrimination across cone-opponent space. If the *adaptive-insensitivity* or *constancy-noise* hypothesis is correct we would expect to find poorer blue-yellow in discrimination of illumination change than discrimination of surface change. If coarser representation explains non-uniform discrimination, as in the *range-accuracy trade-off* hypothesis, we would expect blue-yellow discrimination thresholds to have the same level of asymmetry across conditions.

Some studies have attempted to make this comparison, but the methods used so far have been limited. Weiss and Gegenfurtner (2016) compared illumination discrimination to chromatic detection in twelve colour directions and found similar patterns of relatively poorer discrimination along the daylight locus in both conditions. Pearce (2015) tested illumination discrimination and chromatic detection in four hue directions (red, green, blue, and yellow) in CIE Lu^*v^* and found no difference in chromatic detection thresholds between the four hue directions. Whereas, illumination discrimination was significantly better for green changes than for any other chromatic direction. Pearce interpreted these findings as support for the *adaptive-insensitivity* hypothesis, although there was no difference in illumination discrimination between blue, yellow or red changes.

Current study

This study examines colour discrimination across cone-opponent space, and compares illumination discrimination to surface reflectance discrimination. With this method, we

aim to test support for the *adaptive-insensitivity* hypothesis which predicts relatively poorer discrimination for blue and yellow illuminants than for other colours, but does not predict a similar asymmetry for discrimination of surface reflectance.

To dissociate illumination change from surface reflectance change, we separately manipulated the specular and diffuse components of simulated reflected light in our stimuli. When viewing a surface, the light reaching the eye is a wavelength-by-wavelength product of the spectrum of the illumination (I) and the spectrum of the surface reflectance (R): $I \cdot R$. These two spectra are different components of the light reflected from an object's surface. Specularly reflected light has the spectrum of the illumination (I). The spectrum of diffusely reflected light is product of the illumination spectrum and the object's surface reflectance spectrum ($I \cdot R$). It has been proposed that the varying proportions of diffusely and specularly reflected light reflected from different positions on an object's surface may provide a cue to colour constancy by allowing extrapolation of the chromaticity of the illuminant (D'Zmura & Lennie, 1986; H. C. Lee, 1986; J. N. Yang & Maloney, 2001). Our method of measuring of thresholds for specularly reflected light aims to isolate processes that are responsible for illumination discrimination. As shown in Figure 5.2, when the chromaticities of all the pixels in an image, the colours of the surfaces form a line connecting the object colour to the illumination colour known as "chromaticity convergence " (Hurlbert, 1998). The stimuli in our experiment have three different chromatic surfaces, to aid the observer in chromaticity convergence.

In our study, thresholds for discriminating diffusely reflected light (in the diffuse change condition) were measured without changing specularly reflected light at all. The change in diffusely reflected light was therefore compatible only with a change in the

reflectance spectrum (R) and not with a change in the illumination spectrum (I). This allowed us to estimate thresholds for reflectance change independently from illumination change. In the second (specular change) condition we varied specularly reflected light but held diffusely reflected light constant. This would not occur with a real change in illumination, which affects I both in specularly reflected light (I) and in diffusely reflected light ($I*R$). By varying specularly reflected light and not diffusely reflected light we reduce the ambiguity that is inherent in a physically realistic model over whether illuminant or reflectance have changed. However, we also acknowledge that isolating changes to the specular highlight may create a conflict between cues to the illumination provided by the specular highlight and that provided by the average chromaticity of all three surfaces.

Observers took part in both conditions in a four-alternative force-choice (4AFC) task. In the task, they were asked to spot the “odd-one-out” (target) from three otherwise identical stimuli (distractors). We measured chromatic discrimination thresholds in both conditions, in a scaled version of the MacLeod-Boynton chromaticity diagram (Bosten et al., 2015), in eight different hue directions. This allowed us to test relative discrimination across the chromaticity space, and compare the degree of non-uniformity in colour discrimination between conditions.

5.3 Methods

Participants

Nineteen observers (17 women, mean age = 23.8 years, SD = 4.5) took part. All observers were assessed as having normal colour vision using Ishihara plates (Ishihara, 2010) and the Lanthony Tritan Album (Lanthony, 1998). Participants were paid £7.50 per hour for their participation. Ethical approval was obtained from the Sciences and Technology Cross-Schools Ethical Committee at University of Sussex, and the European Research Council Executive Agency ethics committee.

Design

A four-alternative forced-choice (4AFC) task was used in a repeated measures design. The independent variable was the condition type (diffuse component change or specular component change). The experiment was generally conducted over two 60-minute sessions on different days (average number of days between sessions = 4.1; SD = 6.5). Participants were given as much time as they needed to complete all trials. Two participants did not complete all of the trials in two sessions and so returned for a third session. Participants completed four consecutive blocks of the diffuse component change condition and four consecutive blocks of the specular component change condition per session, resulting in eight blocks of each condition overall. The condition that came first was randomised between participants. There were 64 trials in each block.

Apparatus

A ViSaGe MKII visual stimulus generator in hypercolor mode (which provides chromatic resolution of 14-bits per channel per pixel) drove stimulus presentation (Cambridge Research Systems, Rochester, UK). Stimuli were displayed on a 22-inch

DiamondPlus 2070SB Diamondtron CRT monitor with a resolution of $1,600 \times 1,200$ pixels, and a refresh rate of 100 Hz (Mitsubishi, Tokyo, Japan). The display was gamma-corrected based on the gamma curves of the three primaries, measured with a LS-100 luminance meter (Konica-Minolta, Tokyo, Japan). The colour calibration was based on measurements of the emission spectra of the three primaries made using a SpectraScan PR-655 spectroradiometer (Photo Research Inc., Chatsworth, CA). The experiment was run using custom software written in MATLAB R2016a and the CRS Toolbox extension (The MathWorks Inc., Nantucket, MA).

Stimuli

The stimuli presented in the 4AFC task were three-dimensional spherical, bumpy, glossy “blob” shapes. They were rendered using a model of the interaction between illuminant and material that provided high physical accuracy. The four stimuli presented on each trial appeared in a ‘void’, lit by a single light source. Each blob was ‘patchy’, having three different rendered surface reflectances. The stimuli were generated in a series of steps. The three-dimensional bumpy spherical object was designed using Blender software (<https://www.blender.org/>) (R. J. Lee & Smithson, 2016). The surface geometry of the object was exported to RADIANCE to render high-dynamic-range image files (Ward, 1994; Ward & Shakespeare, 1998). RADIANCE is a package that allows highly accurate simulations of physical objects under different illuminants. It simulates lighting by using ray-tracing techniques to compute radiance values at each pixel in the rendered image.

The blob appeared central in the scene, in a void. The base material specified in RADIANCE for all blobs was “plastic”. The specularity was set to .07 and the roughness was set to .1, which produced a glossy appearance. One directional light

source illuminated the blob. To create each blob stimulus, the Blender object file was rendered as an image file with these specifications using the RADIANCE package (Ward, 1994; Ward & Shakespeare, 1998), run via a custom MATLAB R2016a script (The MathWorks Inc., Massachusetts).

The colours of the stimuli were specified in a scaled version of the MacLeod-Boynton chromaticity diagram (Bosten et al., 2015) based on the Stockman, MacLeod and Johnson (1993) 10-degree cone fundamentals. Sixty-four chromaticities were selected. The colours centred on equal energy white (0.7,1) and varied in eight different hue directions. Colour coordinates were logarithmically spaced, increasing radially from the centre of the plane, so that saturation increased along each colour direction. There were eight coordinates per hue angle. See Figure 5.1 for a representation of the colour coordinates.

The circular points show the colour coordinates of the diffusely reflected light from the variable surface in the diffuse component change condition, and the colour coordinates of the specularly reflected light in the specular component change condition in our version of the MacLeod-Boynton chromaticity diagram. The square black points show the colour coordinates of diffusely reflected light of two surfaces for both conditions. The white circle indicates the colour coordinates of the diffusely reflected light from the third (varying) surface (in the diffuse component change condition) and the specularly reflected light (in the specular component change condition) for the example stimulus presented in Figure 5.2. The centre of the array of chromaticities is metameric with equal-energy white at (0.7, 1), which was the colour coordinates of specularly reflected light in the diffuse component change condition, and of the diffusely reflected light from the third surface in the specular component change condition. For the distractor,

Chapter 5: Why is colour discrimination poorest along the daylight locus?

both the specular component and the third surface were metameric with equal energy white.

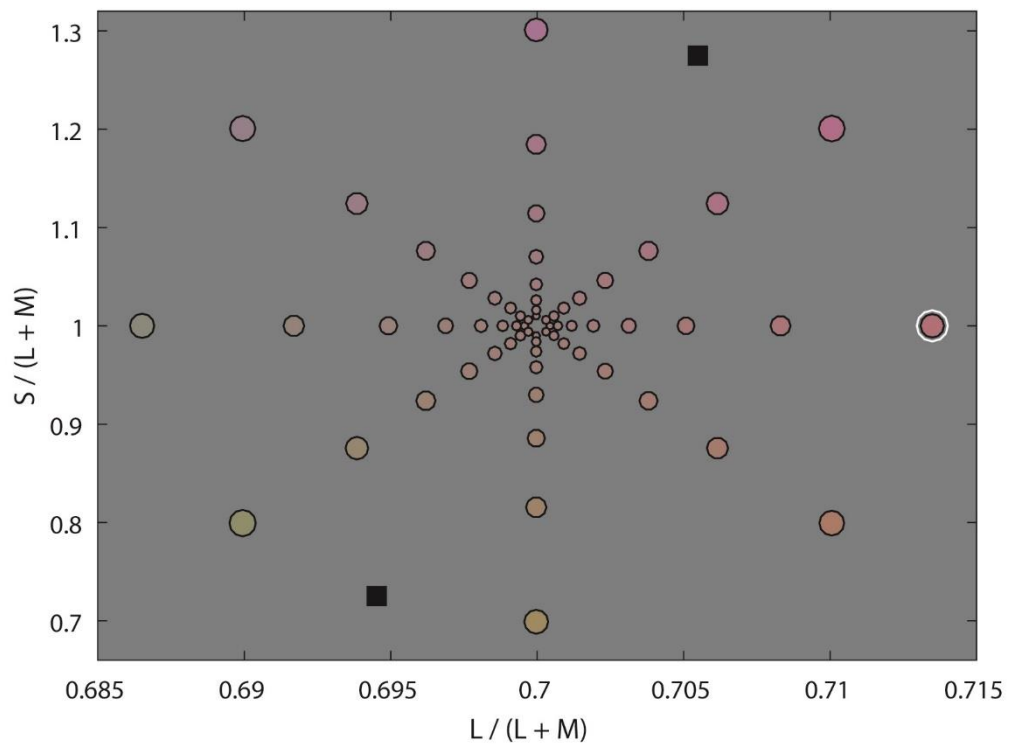


Figure 5.1. Two sets of stimuli were created, one set for the diffuse component change condition and one for the specular component change condition. There were 64 unique stimuli per set, plus one “distractor” stimulus that was the same in both conditions. The distractor stimulus comprised three of the four stimuli displayed on every trial in the 4AFC task. For the distractor stimulus, the chromaticity of the specularly reflected light was metameric with equal energy white (0.7, 1 in MacLeod-Boynton chromaticity diagram). Two of the surfaces had a diffuse component indicated by the square black points in Figure 5.2 and the third surface had a diffuse component metameric with equal energy white.

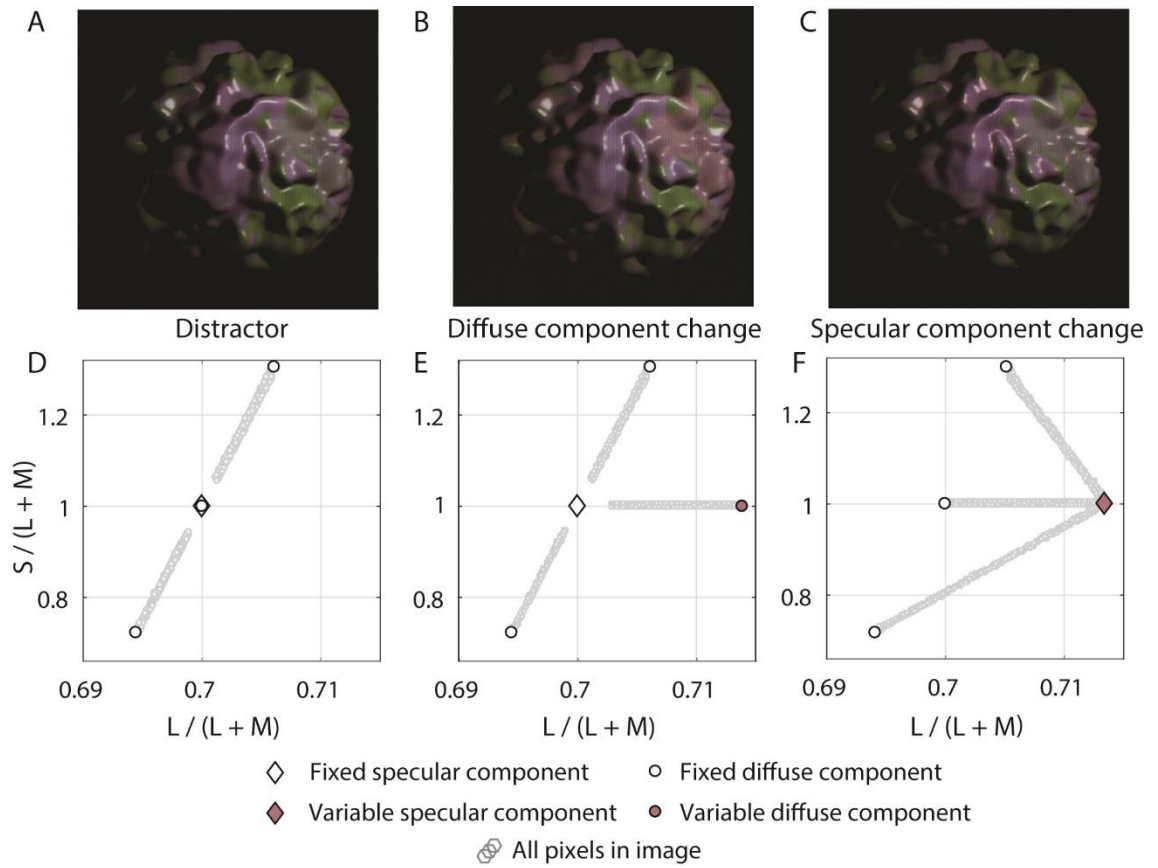


Figure 5.2. The top panels show: A) the distractor stimulus, B) an example stimulus from the diffuse component change condition and, C) an example stimulus from the specular component change condition. The lower panels (D-F) show the chromaticity distributions from the above example stimuli.

For the diffuse component change condition, the chromaticity of the specularly reflected light was always metameric with equal energy white (0.7, 1 in the MacLeod-Boynton chromaticity diagram), and two of the three surfaces had the same diffuse component in each of the 64 stimuli (black squares in Figure 5.1). The third surface of the target stimulus was different in each of the 64 stimuli in this condition, corresponding to the 64 colour coordinates shown in Figure 5.1. Panel B of Figure 5.2 shows an example of a diffuse component change stimulus. The third (variable) surface is a pinkish colour, and can be most clearly seen in the middle-right area of the blob. The two surfaces with fixed diffuse components are purplish and greenish.

For the specular component change condition, the specularly reflected light varied across the 64 colours shown in Figure 5.1, whilst the diffuse components stayed constant. The diffuse components for two of the surfaces had chromaticities shown by the black squares in Figure 5.1 (as for the diffuse component change stimuli and the distractor). The diffuse component for the third surface was metameric with equal energy white (as for the distractor). Panel C of Figure 5.2 shows an example of one stimulus in the specular change condition. The difference between this stimulus and the distractor (panel A) can be seen (for example) by looking at the specular highlight near the top-left of the blob. The distractor reflects a whitish light in its specular highlight whereas the specular component change stimulus reflects a pinkish light. The distributions of chromaticities present in each of the example stimuli can be seen in the corresponding lower panels (D-F) of Figure 5.2. The plots show that there is a gradient of chromaticities between the body colour (diffusely reflected light) and the specular highlight (D'Zmura & Lennie, 1986; Lee, 1986), formed from different ratios of mixture of the two components. Panel D shows that the distractor has two different body colours (the third body colour is the same as the specular component, but with a

much lower intensity), so there are two chromaticity gradients, between the specular highlight (0.7, 1) and the chromaticities of the diffuse components (indicated by black squares in Figure 5.1). Panel E shows that the diffuse component change example stimulus has three surfaces with body colours different from the specular component. Panel F shows that the specular component change condition example has chromaticity gradients that converge on the varying chromaticity of the illuminant.

Procedure

On each trial of the experiment, four stimuli were presented, one in each quadrant of the screen. Three of the stimuli were identical (i.e. the distractor stimulus) and one stimulus was different (the target). The target stimulus was controlled using the method of constant stimuli procedure. In the diffuse component change condition, one stimulus had a different diffuse component chromaticity (with 64 different values, see Figure 5.1), for one of the stimulus surfaces. In the specular component change condition, one stimulus had a different specular component chromaticity (with the same 64 different values as for the diffuse component change condition). A black background was used throughout the experiment. Participants were asked to determine which of the four stimuli the odd one out was by pressing a corresponding key on a keyboard. They received auditory feedback: a low tone (400hz for 100ms) indicated an incorrect response and a high tone (1000hz for 100ms) indicated a correct response. Participants were asked to look carefully at each stimulus and to take as long as they needed to respond.

The experiment began with a training phase, which consisted of trials showing the maximally saturated target stimulus, drawn randomly from the eight hue angles. The experiment proceeded to the test phase when participants completed six consecutive

training trials correctly. The experiment took place in a dark room and participants sat at a distance of 50cm from the monitor (not constrained). Each blob stimulus was 10cm x 10cm on the monitor, corresponding to a visual angle of 11° .

5.4 Data Analysis

Our aim was to determine whether participants' discrimination for colours in the blue-yellow direction was relatively poorer for the specular component, compared to the diffuse component. Proportion correct data for each participant from the specular and diffuse component change conditions were fit parametrically with logistic functions from the Modelfree toolbox for MATLAB (Zchaluk & Foster, 2009). Each condition had eight different hue angles, and each angle had eight stimulus levels, increasing logarithmically in saturation (see Figure 5.1 for stimulus coordinates). Example plots showing proportion correct and the fitted psychometric functions for one participant can be seen in Figure 5.3.

A discrimination threshold was taken from each psychometric function at the 60% correct level. Discrimination thresholds were taken for each of the eight hue angles in both conditions. One participant could not have their threshold measured at the 60% level as they were at ceiling for all hue angles in the specular condition, therefore they were excluded. Three further participants could not have their threshold measured because of poor performance in the specular condition for some hue angles, leading to a floor effect within the available gamut. These participants were also excluded. Therefore, the analysis is based on data from 15 participants.

The discrimination thresholds were then plotted in our version of the MacLeod-Boynton chromaticity diagram and a best-fitting ellipse was fit through each set of eight thresholds for every participant, for each of the two conditions. An example from one participant of thresholds and the fitted ellipses can be seen in Figure 5.4. The thresholds were then transformed into a normalised space, which equates the variability in the thresholds (within each participant) along the x and y axes. This calculation is done

post-hoc; any bias remaining after the standardisation tells us about differences between the conditions independent of individual differences in the LMS sensitivity. This is done because the scaling of S relative to L and M cones in the MacLeod-Boynton diagram is arbitrary, but we wish to equalise the scaling to test the uniformity of colour discrimination performance in other hue directions once the scaling differences of the two cardinal axes are factored out (Bosten et al., 2015).

An ‘axis ratio’ was then calculated for each ellipse in the normalised space. In the normalised space, the major axis of the ellipse will always be along either the negative diagonal or the positive diagonal. The axis ratio measures the ratio of the length of the negatively sloping axis to the length of the positively sloping axis. An axis ratio of 1 indicates equal discrimination performance in each hue direction (i.e. a circle). An axis ratio greater than 1 indicates reduced performance along the negative diagonal relative to performance along the positive diagonal, which indicates relatively poor performance along the blue-yellow axis.

5.5 Results

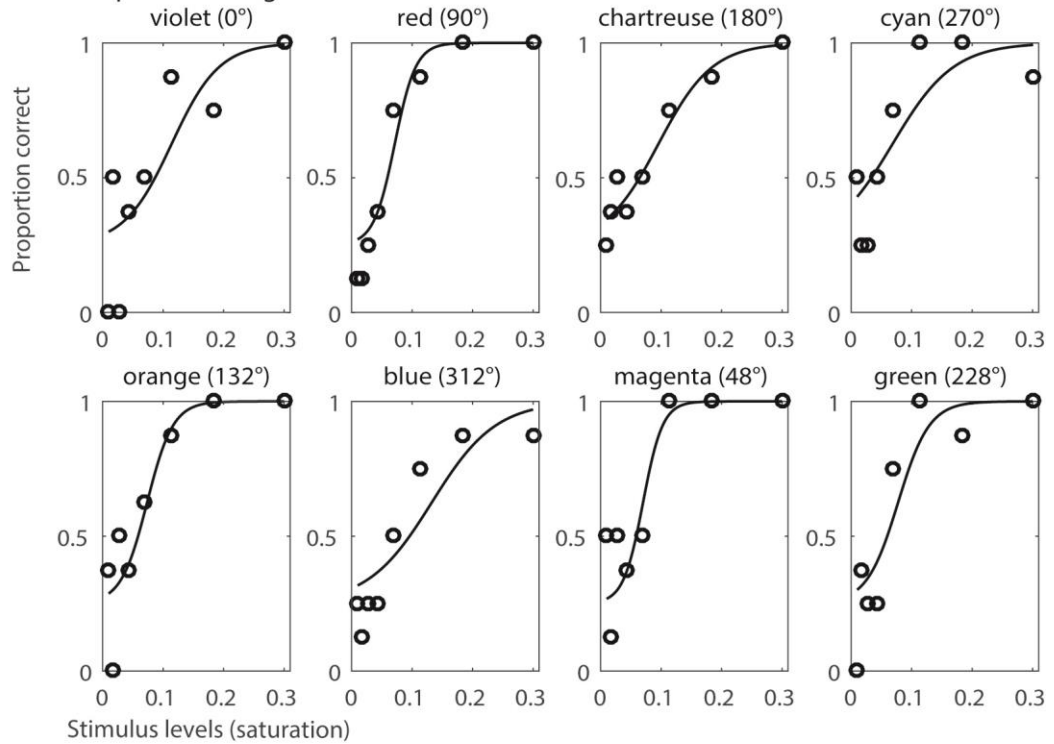
The results are based on data from 15 participants. Mean axis ratios for the diffuse component discrimination ellipse (mean = 1.35; SEM = 0.12) and the specular component discrimination ellipse (mean = 1.65; SEM = 0.11) are shown in Figure 5.5. Single sample *t*-tests were conducted on the axis ratios from each condition, which showed that the mean axis ratio in both conditions was significantly different from 1 (diffuse $t(14) = 11.62, p < .001$; specular $t(14) = 15.29, p < .001$). This indicates that participants were significantly worse in the direction of the negative diagonal (roughly blue-yellow) than the orthogonal (roughly red-green) direction when making colour discrimination judgements for both a diffuse component change and a specular component change. Furthermore, confidence intervals of the thresholds were calculated, and variance in the diffuse component change condition was compared to variance in the specular component change condition. Confidence intervals were collapsed across hue directions to get an average confidence interval for each condition per participant. A paired samples *t*-test revealed that there was greater variance in the specular component change condition (mean = 0.09) to the diffuse component change condition (mean = 0.06; $t(14) = 2.93, p = 0.011$). This was to be expected as the specular component change condition was more challenging.

Next, our experimental hypothesis was tested. That is, whether there is a significant difference between the axis ratios for the two conditions, indicating that participants are relatively poorer at discriminating in the direction of the negative diagonal when the specular component changes, compared to when the diffuse component changes. A paired samples *t*-test indicated no significant difference between the mean axis ratios of the two conditions ($t(14) = -1.96, p = 0.070$). The data show a trend for relatively poorer performance at discriminating along the negative diagonal when the specular

component changes compared to when the diffuse component changes, but the difference is not significant.

There was no evidence of a speed-accuracy trade-off in the response data. Correlational analyses revealed no significant correlation between participants' thresholds (averaged across hue angles) and their average reaction times in the specular component change condition ($r = -0.468$, $p = .078$), nor in the diffuse component change condition ($r = -0.372$, $p = .172$). There was also no significant correlation between participants' axis ratios in the diffuse and specular component change conditions ($r = 0.313$, $p = .257$).

A. Diffuse component change



B. Specular component change

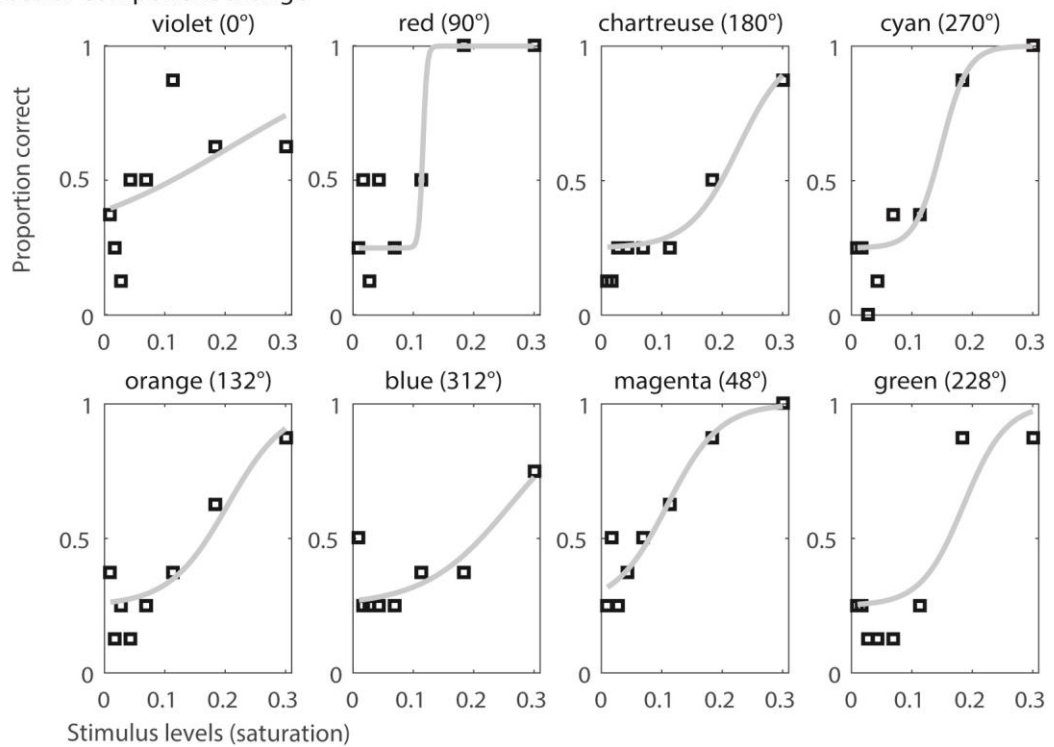


Figure 5.3. Example results from one participant for, A) the diffuse component change condition, and B) the specular component change condition. The points show the average proportion correct for each hue angle (labelled above each plot; circular points

for the diffuse condition and square points for the specular condition). Each hue angle has eight stimulus levels, increasing logarithmically in saturation (x-axis). The stimulus levels on the x-axis denote increasing colour saturation. Saturation (s) and hue angle (θ) can be converted to MacLeod Boynton chromaticity coordinates (in our diagram based on the Stockman and MacLeod and Johnson (1993) cone fundamentals) via the following two equations: (i) $S/(L+M) = 1.0 + s \cdot \cos(\theta)$; (ii) $L/(L+M) = 0.7 + 0.045 \cdot s \cdot \cos(\theta)$. The curves show logistic psychometric functions fit to the data. Thresholds were taken from each curve at the 0.6 level.

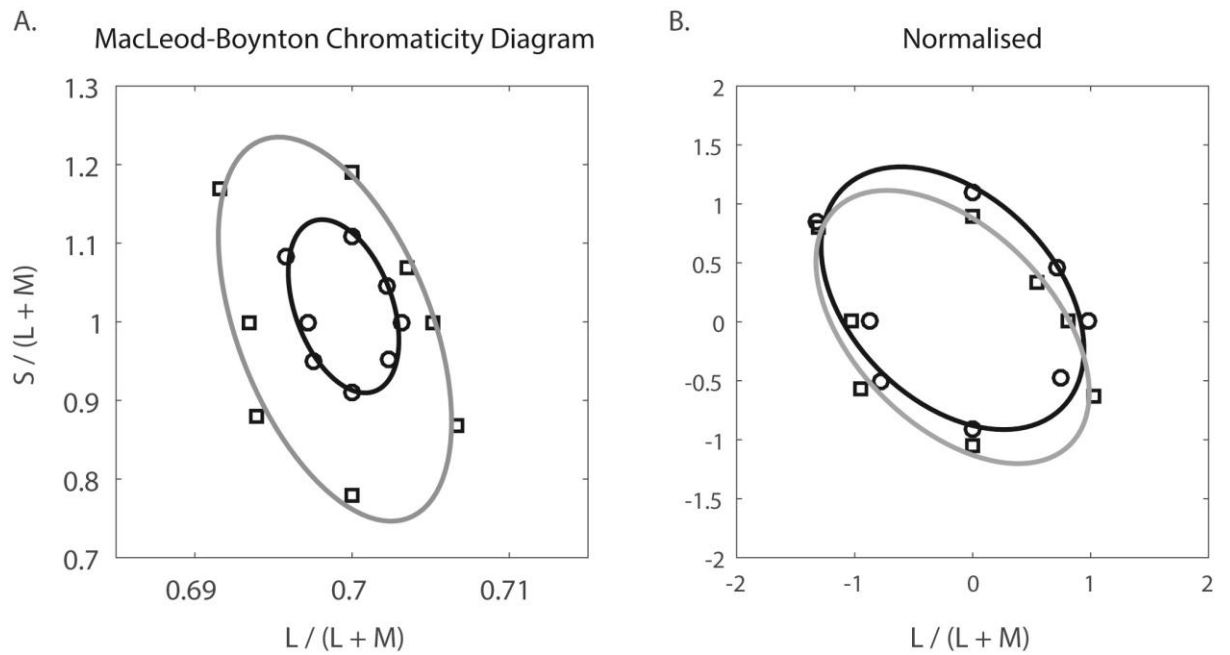


Figure 5.4. Threshold data taken from the psychometric functions shown in Figure 5.3.

The circular points show the diffuse component discrimination thresholds and the square points show the specular component discrimination thresholds. A black ellipse is fit to the diffuse component thresholds and a grey ellipse is fit to the specular component thresholds. Panel A shows the thresholds in our version of the MacLeod-Boynton chromaticity diagram and panel B shows the same data in a normalized space, where variability along the x and y axes is equated.

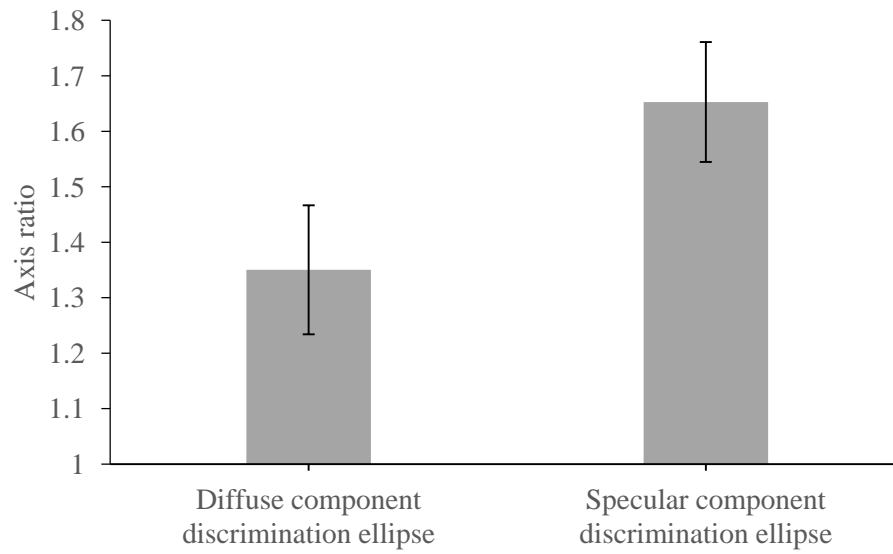


Figure 5.5. The mean axis ratios for the diffuse component discrimination ellipse and the specular component discrimination ellipse, averaged across participants ($N = 15$). Error bars = 1 SEM. There is no significant difference in the axis ratios between the conditions.

5.6 Discussion

The aim of our study was to test the adaptive insensitivity hypothesis of asymmetric chromatic discrimination in a cone-opponent chromaticity diagram. This hypothesis states that discrimination is relatively poorer along the blue-yellow direction (i.e. the daylight locus), compared to the orthogonal direction, as an aid to colour constancy (Pearce et al., 2014; Radonjic, Pearce, et al., 2016). Our study compared discrimination of a specular component change (a cue to the illumination) with discrimination of a diffuse component change (a cue to surface reflectance). If the adaptive insensitivity hypothesis were correct, we would expect relatively poorer blue-yellow discrimination only for changes in the specular component and not for changes in the diffuse component. Because, if relatively poorer discrimination along the blue-yellow direction were an aid to colour constancy, this would only be adaptive for illumination discrimination, not for surface reflectance discrimination.

In the current study, we separately manipulated the specular and diffuse components of light reflected in rendered blob stimuli. Participants took part in a 4AFC task, with two conditions, where they were asked to detect changes in a) the specular component and b) the diffuse component of light reflected from the blob stimuli. In each condition, the specular and diffuse components of simulated reflected light varied in eight different hue directions, and had eight different levels of saturation in a scaled version of MacLeod-Boynton chromaticity diagram (Bosten et al., 2015). We quantified asymmetry in discrimination for a specular component change, compared to that for a diffuse component change, by plotting the discrimination thresholds for hue angle from each condition in a normalised space, and then fitting an ellipse to these points. We then calculated axis ratios (the ratio of the length of the negative diagonal to the length of the positive diagonal) for the discrimination ellipses that resulted from both conditions.

Axis ratios larger than one indicate relatively poorer discrimination in the blue-yellow direction (Bosten et al., 2015).

In both the specular component and diffuse component conditions, the axis ratios of the colour discrimination ellipses were significantly greater than one, meaning that colour discrimination was poorer in the blue-yellow direction than in the red-green direction. The adaptive insensitivity hypothesis predicts that discrimination thresholds would be raised in the blue-yellow direction only for the specular component, whereas discrimination thresholds for the diffuse component would be uniform across the normalised space. Therefore, the adaptive insensitivity hypothesis in its pure form is not supported by our findings.

The range-accuracy trade-off hypothesis states that relatively poorer blue-yellow discrimination is due to coarse representation in the visual system, as there is greater variation in this direction in natural scene statistics (Bosten et al., 2015; MacLeod, 2003; Macleod & von der Twier, 2001). This hypothesis predicts relatively poorer discrimination in the blue-yellow direction for both conditions, with no difference in the extent of the asymmetry. In the current study, there was no significant difference in the axis ratios between the two conditions. This finding is in line with the prediction from the range-accuracy trade-off hypothesis. However, there was a strong but non-significant trend towards greater elongation along the blue-yellow direction of the discrimination ellipse for the specular component, compared to the discrimination ellipse for the diffuse component. Which theory best explains this pattern of results?

The observed trend suggests that the asymmetry in discrimination across cone-opponent space may be related to colour constancy. We propose that uncertainty about the illumination may be a contributing factor to this phenomenon, but does not fully explain

asymmetric discrimination, or the axis ratio of the discrimination ellipse for the diffuse component would not be significantly greater than one.

The theories discussed so far that seek to explain the observed asymmetry in colour discrimination all rely on cortical mechanisms. However, it has also been suggested that there may be a low-level, retinal channel (e.g. a class of retinal ganglion cell) tuned orthogonally to the blue-yellow axis (i.e. roughly along red-green axis), which boosts colour discrimination performance along that axis (Danilova & Mollon, 2012a, 2012b, 2014). However, any account of asymmetric colour discrimination that relies on a retinal mechanism would not be compatible with a difference in axis ratios between our experimental conditions: it does not account for the trend that we observe towards relatively poorer blue-yellow discrimination for a colour change in specularly reflected light than for a colour change in diffusely reflected light.

Conclusion

Our results show that discrimination of a colour change in the diffuse component and in the specular component of the light reflected from an object is relatively poorer in the blue-yellow colour direction, than in the orthogonal direction. Furthermore, we find a (non-significant) trend that suggests that this asymmetry in colour discrimination is greater for specularly reflected light (a cue to illumination) than for diffusely reflected light (a cue to reflectance). Our data are most consistent with the idea that a range-accuracy trade-off contributes to asymmetric colour discrimination in cone-opponent colour space. The distribution of chromaticities in natural scenes contains greatest variance in the blue-yellow direction: our visual system must therefore sacrifice accuracy for range. Our finding that colour discrimination performance is biased in a blue-yellow direction both for specularly reflected light and for diffusely reflected light

supports this. However, there is a strong trend in our data for relatively poorer blue-yellow discrimination for specularly reflected light than for diffusely reflected light. This trend provides some support for the adaptive insensitivity hypothesis (that poor discrimination along the daylight locus is adaptive for colour constancy) or the constancy noise hypothesis (that colour constancy corrections, usually along the daylight axis, introduce noise into colour judgements).

There is still uncertainty about the origin of the asymmetric pattern of colour discrimination performance in cone opponent space. Is the asymmetry tuned to the chromatic statistics of the natural environment? If so, does the tuning occur over lifetime, or is it evolved? Further studies could investigate these questions through cross-cultural and developmental research. For example, a study of the development of chromatic discrimination in infants and young children could reveal the extent to which the asymmetry is innate.

Further work is also needed to untangle the different theories of asymmetric colour discrimination across cone-opponent space: the adaptive insensitivity hypothesis, the constancy noise hypothesis and the range-accuracy trade-off hypothesis. The adaptive insensitivity hypothesis predicts that both the colour constancy index and the variance of colour constancy settings would be relatively greater along the blue-yellow axis. The constancy noise hypothesis predicts that the variance of settings would be relatively greater along a blue-yellow axis than other colour axes, but not the constancy index. Results from studies such as Brainard (1998) and Delahunt and Brainard (2004) are discouraging for the adaptive insensitivity hypothesis, but an ideal experiment would measure both index and variance. The three theories of asymmetry in colour

discrimination are not mutually exclusive, and one or more may be contributing to the observed findings in this study: more work needs to be done to determine this.

References

- Abbatecola, C., Gerardin, P., Knoblauch, K., & Kennedy, H. (2016). Face and voice contributions to gender discrimination. *Perception ECVP*, 45, 331.
- Abdi, H. (2009). Signal Detection theory (SDT). In B. McGaw, P. L. Peterson, & E. Baker (Eds.), *Encyclopedia of Education* (3rd editio). New York: Elsevier.
- Abney, de W. W. (1909). On the Change in Hue of Spectrum Colours by Dilution with White Light. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 83(560), 120–127. <http://doi.org/10.1098/rspa.1909.0085>
- Abramov, I., Gordon, J., Hendrickson, A., Hainline, L., Dobson, V., & LaBossiere, E. (1982). The retina of the newborn human infant. *Science*, 217(4556), 265–7.
- Abrams, A. B., Hillis, J. M., & Brainard, D. H. (2007). The relation between color discrimination and color constancy: when is optimal adaptation task dependent? *Neural Computation*, 19(10), 2610–2637. <http://doi.org/10.1162/neco.2007.19.10.2610>
- Agarwal, V., Gribok, A. V., & Abidi, M. A. (2007). Machine learning approach to color constancy. *Neural Networks*, 20(5), 559–563. <http://doi.org/10.1016/j.neunet.2007.02.004>
- Albert, A., & Anderson, J. A. (1984). On the Existence of Maximum Likelihood Estimates in Logistic Regression Models. *Biometrika*, 71(1), 1. <http://doi.org/10.2307/2336390>
- Allen, D., Banks, M. S., & Norcia, A. M. (1993). Does chromatic sensitivity develop more slowly than luminance sensitivity? *Vision Research*, 33(17), 2553–62.

References

- Allen, E. C., Beilock, S. L., & Shevell, S. K. (2011). Working memory is related to perceptual processing: A case from color perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 1014–1021.
<http://doi.org/10.1037/a0023257>
- Allen, E. C., Beilock, S. L., & Shevell, S. K. (2012). Individual differences in simultaneous color constancy are related to working memory. *Journal of the Optical Society of America A*, 29(2), A52.
<http://doi.org/10.1364/JOSAA.29.000A52>
- Allred, S. R., & Olkkonen, M. (2015). The effect of memory and context changes on color matches to real objects. *Attention, Perception, & Psychophysics*.
<http://doi.org/10.3758/s13414-014-0810-4>
- Álvaro, L., Linhares, J. M. M., Moreira, H., Lillo, J., & Nascimento, S. M. C. (2017). Robust colour constancy in red-green dichromats. *PLOS ONE*, 12(6), e0180310.
<http://doi.org/10.1371/journal.pone.0180310>
- Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America. A, Optics and Image Science*, 3(10), 1743–51.
- Ashby, F. G. (2007). Multivariate Probability Distributions. In *Risk and Reliability Analysis* (pp. 272–312). Reston, VA: American Society of Civil Engineers.
<http://doi.org/10.1061/9780784408919.ch07>
- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10(1), 48–53.
<http://doi.org/10.1111/j.1467-7687.2007.00563.x>
- Banks, M. S., & Bennett, P. J. (1988). Optical and photoreceptor immaturities limit the

References

- spatial and chromatic vision of human neonates. *Journal of the Optical Society of America A*, 5(12), 2059. <http://doi.org/10.1364/JOSAA.5.002059>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.
- Beer, R. D., Dinca, A., & MacLeod, D. I. A. (2010). Ideal white can be yellowish or bluish, but not reddish or greenish. *Journal of Vision*, 6(6), 417–417. <http://doi.org/10.1167/6.6.417>
- Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2018). Facial color is an efficient mechanism to visually transmit emotion. *Proceedings of the National Academy of Sciences*, 115(14), 3581–3586. <http://doi.org/10.1073/pnas.1716084115>
- Bimler, D. L. (2011). Universal trends and specific deviations: Multidimensional scaling of colour terms from the World Color Survey. In *New Directions in Colour Studies* (pp. 13–26). Amsterdam: John Benjamins Publishing Company. <http://doi.org/10.1075/z.167.05bim>
- Bimler, D. L., Kirkland, J., & Jameson, K. A. (2004). Quantifying variations in personal color spaces: Are there sex differences in color vision? *Color Research & Application*, 29(2), 128–134. <http://doi.org/10.1002/col.10232>
- Bird, C. M., Berens, S. C., Horner, A. J., & Franklin, A. (2014). Categorical encoding of color in the brain. *Proceedings of the National Academy of Sciences*, 111(12), 4590–4595. <http://doi.org/10.1073/pnas.1315275111>
- Bornstein, M. H. (1976). Infants Are Trichromats. *Journal of Experimental Child*

References

- Psychology*, 21, 425–445. [http://doi.org/https://doi.org/10.1016/0022-0965\(76\)90072-2](http://doi.org/https://doi.org/10.1016/0022-0965(76)90072-2)
- Bornstein, M. H. (1985a). Colour-name versus shape-name learning in young children. *Journal of Child Language*, 12(02), 387–393. <http://doi.org/10.1017/S0305000900006498>
- Bornstein, M. H. (1985b). On the development of color naming in young children: Data and theory. *Brain and Language*, 26(1), 72–93. [http://doi.org/10.1016/0093-934X\(85\)90029-X](http://doi.org/10.1016/0093-934X(85)90029-X)
- Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology. Human Perception and Performance*, 2(1), 115–129.
- Bosten, J. M., Beer, R. D., & MacLeod, D. I. A. (2015). What is white? *Journal of Vision*, 15(16), 5. <http://doi.org/10.1167/15.16.5>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *JOSA A*. <http://doi.org/0740-3232/98/020307-19>
- Brainard, D. H. (2009). Color constancy. In B. Goldstein (Ed.), *Sage Encyclopedia of Perception* (pp. 253–257). Los Angeles: SAGE Publications.
- Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). *Color constancy in the nearly natural image. I. Asymmetric matches*.
- Brainard, D. H., Longere, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B.

References

- (2006). Bayesian model of human color constancy. *Journal of Vision*, 6(11), 10–10. <http://doi.org/10.1167/6.11.10>
- Brainard, D. H., Wandell, B. A., & Chichilnisky, E.-J. (1993). Color Constancy: From Physics to Appearance. *Current Directions in Psychological Science*, 2(5), 165–170.
- Brown, A. M., & Lindsey, D. T. (2013). Infant color vision and color preferences: a tribute to Davida Teller. *Visual Neuroscience*, 30(5–6), 243–50. <http://doi.org/10.1017/S0952523813000114>
- Brown, A. M., Lindsey, D. T., & Guckes, K. M. (2011). Color names, color categories, and color-cued visual search: Sometimes, color perception is not categorical. *Journal of Vision*, 11(12), 1–21. <http://doi.org/10.1167/11.12.2>
- Brown, A. M., & Teller, D. Y. (1989). Chromatic opponency in 3-month-old human infants. *Vision Research*, 29(1), 37–45.
- Brown, R. O. (1994). Backgrounds and Illuminants: The Yin and Yang of Colour Constancy. In *Colour Perception: Mind and the physical world* (pp. 247–278). Oxford University Press. <http://doi.org/10.1093/acprof:oso/9780198505006.003.0008>
- Buchsbaum, G., & Gottschalk, A. (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 220(1218), 89–113.
- Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: the representation of hue, saturation, and brightness.

References

- Perception & Psychophysics*, 43(5), 494–507. <http://doi.org/10.3758/BF03207885>
- Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*. Chicago: University of Chicago Press.
- Changizi, M. A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology Letters*, 2(2), 217–221.
<http://doi.org/10.1098/rsbl.2006.0440>
- Chauhan, T., Perales, E., Xiao, K., Hird, E., Karatzas, D., & Wuerger, S. (2014). The achromatic locus: Effect of navigation direction in color space. *Journal of Vision*, 14(1), 25–25. <http://doi.org/10.1167/14.1.25>
- Chien, S. H.-L., Palmer, J., & Teller, D. Y. (2005). Achromatic contrast effects in infants: Adults and 4-month-old infants show similar deviations from Wallach's ratio rule. *Vision Research*, 45(22), 2854–2861.
<http://doi.org/10.1016/j.visres.2004.07.035>
- Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America*, 86(23), 9631–5.
- Civan, A., Teller, D. Y., & Palmer, J. (2005). Relations Among Spontaneous Preferences, Familiarized Preferences, and Novelty Effects: Measurements With Forced-Choice Techniques. *Infancy*, 7(2), 111–142.
http://doi.org/10.1207/s15327078in0702_1
- Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., Özgen, E., & Davies, I. R. L. (2012). Neural correlates of acquired color category effects. *Brain and Cognition*,

References

- 80(1), 126–43. <http://doi.org/10.1016/j.bandc.2012.04.011>
- Cohen, D. J. (2003). Direct estimation of multidimensional perceptual distributions: Assessing hue and form. *Perception & Psychophysics*, 65(7), 1145–1160. <http://doi.org/10.3758/BF03194841>
- Commission Internationale de L'Eclairage (CIE). International Lighting Vocabulary - 2nd edition. CIE DIS 017/E:2016 ILV. (2017).
- Conway, B. R. (2009). Color vision, cones, and color-coding in the cortex. *Neuroscientist*, 15(3), 274–290. <http://doi.org/10.1177/1073858408331369>
- Conway, B. R., Chatterjee, S., Field, G. D., Horwitz, G. D., Johnson, E. N., Koida, K., & Mancuso, K. (2010). Advances in Color Science: From Retina to Behavior. *Journal of Neuroscience*, 30(45), 14955–14963. <http://doi.org/10.1523/JNEUROSCI.4348-10.2010>
- Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: an analysis of viewing strategies. *Vision Research*, 35(17), 2431–48.
- D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A*, 3(10), 1662. <http://doi.org/10.1364/JOSAA.3.001662>
- Danilova, M. V., & Mollon, J. D. (2012a). Cardinal axes are not independent in color discrimination. *Journal of the Optical Society of America A*, 29(2), A157. <http://doi.org/10.1364/JOSAA.29.00A157>
- Danilova, M. V., & Mollon, J. D. (2012b). Foveal color perception: Minimal thresholds at a boundary between perceptual categories. *Vision Research*, 62, 162–172.

References

<http://doi.org/10.1016/j.visres.2012.04.006>

Danilova, M. V., & Mollon, J. D. (2014). Symmetries and asymmetries in chromatic discrimination. *Journal of the Optical Society of America A*, 31(4), A247.

<http://doi.org/10.1364/JOSAA.31.00A247>

Dannemiller, J. L. (1989). A test of color constancy in 9- and 20-week-old human infants following simulated illuminant changes. *Developmental Psychology*, 25(2), 171–184. <http://doi.org/10.1037//0012-1649.25.2.171>

Dannemiller, J. L., & Hanko, S. A. (1987). A test of color constancy in 4-month-old human infants. *Journal of Experimental Child Psychology*, 44(2), 255–67.

Darwin, C. H. (1877). A biographical sketch of a young infant. *Kosmos*, (1), 367–376.

Datasheet: HP DesignJet Z3200 Photo Printer series. (2017). Retrieved from <http://www8.hp.com/h20195/V2/GetPDF.aspx/4aa4-4150eee>

Daugirdiene, A., Kulikowski, J. J., Murray, I. J., & Kelly, J. M. F. (2016). Test illuminant location with respect to the Planckian locus affects chromaticity shifts of real Munsell chips. *Journal of the Optical Society of America A*, 33(3), A77. <http://doi.org/10.1364/JOSAA.33.000A77>

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, 4(2), 1. <http://doi.org/10.1167/4.2.1>

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, 357, 241–65.

References

- Devinck, F., & Knoblauch, K. (2012). A common signal detection model accounts for both perception and discrimination of the watercolor effect. *Journal of Vision*, 12, 19–19. <http://doi.org/10.1167/12.3.19>
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5(July), 781. <http://doi.org/10.3389/fpsyg.2014.00781>
- Diesendruck, G., & Bloom, P. (2003). How Specific is the Shape Bias? *Child Development*, 74(1), 168–178. <http://doi.org/10.1111/1467-8624.00528>
- Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on creator's intent in extending names for Artifacts. *Psychological Science*, 14(2), 164–169. <http://doi.org/10.1111/1467-9280.t01-1-01436>
- Dobkins, K. R. (2009). Does Visual Modularity Increase Over the Course of Development? *Optometry and Vision Science*, 86(6), E583–E588. <http://doi.org/10.1097/OPX.0b013e3181a72854>
- Dobkins, K. R., & Anderson, C. M. (2002). Color-Based Motion Processing Is Stronger in Infants Than in Adults. *Psychological Science*, 13(1), 76–80. <http://doi.org/10.1111/1467-9280.00414>
- Dobkins, K. R., Anderson, C. M., & Lia, B. (1999). Infant temporal contrast sensitivity functions (tCSFs) mature earlier for luminance than for chromatic stimuli: evidence for precocious magnocellular development? *Vision Research*, 39(19), 3223–3239. [http://doi.org/10.1016/S0042-6989\(99\)00020-6](http://doi.org/10.1016/S0042-6989(99)00020-6)
- Dominy, N. J., & Lucas, P. W. (2001). Ecological importance of trichromatic vision to primates. *Nature*, 410(6826), 363–366. <http://doi.org/10.1038/35066567>

References

- Ekman, G. (1954). Dimensions of Color Vision. *The Journal of Psychology*, 38(2), 467–474. <http://doi.org/10.1080/00223980.1954.9712953>
- Ekroll, V. (2005). *On the nature of simultaneous colour contrast*. Christian-Albrechts-Universität zu Kiel.
- Falmagne, J. (2002). *Elements of psychophysical theory*. Oxford: Oxford University Press.
- Foster, D. H. (2003). Does colour constancy exist? *Trends in Cognitive Sciences*, 7(10), 439–443. <http://doi.org/10.1016/j.tics.2003.08.002>
- Foster, D. H. (2011). Color constancy. *Vision Research*, 51(7), 674–700. <http://doi.org/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. <http://doi.org/10.1017/S0952523806233455>.Color
- Foster, D. H., Amano, K., Nascimento, S. M. C., & Foster, M. J. (2006). Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23(10), 2359. <http://doi.org/10.1364/JOSAA.23.002359>
- Franklin, A. (2006). Constraints on children's color term acquisition. *Journal of Experimental Child Psychology*, 94(4), 322–327. <http://doi.org/10.1016/j.jecp.2006.02.003>
- Franklin, A., Bevis, L., Ling, Y., & Hurlbert, A. C. (2010). Biological components of colour preference in infancy. *Developmental Science*, 13(2), 346–354. <http://doi.org/10.1111/j.1467-7687.2009.00884.x>

References

- Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, 90(2), 114–141. <http://doi.org/10.1016/j.jecp.2004.10.001>
- Franklin, A., & Davies, I. R. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology*, 22(3), 349–377.
- Franklin, A., Pitchford, N., Hart, L., Davies, I. R. L., Clausse, S., & Jennings, S. (2008). Salience of primary and secondary colours in infancy. *British Journal of Developmental Psychology*, 26(4), 471–483. <http://doi.org/10.1348/026151007X256672>
- Garner, W. R. (1974). *The processing of information and structure*. Hillside, NJ: Erlbaum.
- Garrigan, P., & Kellman, P. J. (2008). Perceptual learning depends on perceptual constancy. *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), 2248–53. <http://doi.org/10.1073/pnas.0711878105>
- Gegenfurtner, K. R. (2003). Sensory systems: Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, 4(7), 563–572. <http://doi.org/10.1038/nrn1138>
- Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience*, 13(1), 161–72.
- Gerardin, P., Devinck, F., Dojat, M., & Knoblauch, K. (2014). Contributions of contour frequency, amplitude, and luminance to the watercolor effect estimated by conjoint measurement. *Journal of Vision*, 14(4), 9–9. <http://doi.org/10.1167/14.4.9>
- Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination.

References

- Journal of Experimental Psychology: General*, 123(2), 178–200.
<http://doi.org/10.1037//0096-3445.123.2.178>
- Golz, J., & MacLeod, D. I. A. (2002). Influence of scene statistics on colour constancy. *Nature*, 415(6872), 637–640. <http://doi.org/10.1038/415637a>
- Gouko, M., & Kobayashi, Y. (2010). State Representation with Perceptual Constancy Based on Active Motion (pp. 100–109). http://doi.org/10.1007/978-3-642-17248-9_11
- Granrud, C. (2006). Size constancy in infants: 4-month-olds' responses to physical versus retinal image size. *Journal of Experimental Psychology. Human Perception and Performance*, 32(6), 1398–404. <http://doi.org/10.1037/0096-1523.32.6.1398>
- Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Guild, J. (1932). The Colorimetric Properties of the Spectrum. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 230(681–693), 149–187. <http://doi.org/10.1098/rsta.1932.0005>
- Hammarrenger, B., Leporé, F., Lippé, S., Labrosse, M., Guillemot, J. P., & Roy, M. S. (2003). Magnocellular and parvocellular developmental course in infants during the first year of life. *Documenta Ophthalmologica*, 107(3), 225–233.
<http://doi.org/10.1023/B:DOOP.0000005331.66114.05>
- Hansen, T., Giesel, M., & Gegenfurtner, K. R. (2008). Chromatic discrimination of natural objects. *Journal of Vision*, 8(1), 2. <http://doi.org/10.1167/8.1.2>
- Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal

References

- context on color categories and color constancy. *Journal of Vision*, 7(4), 2–2.
<http://doi.org/10.1167/7.4.2>
- Hansmann-Roth, S., & Mamassian, P. (2017). A Glossy Simultaneous Contrast: Conjoint Measurements of Gloss and Lightness. *I-Perception*, 8(1), 204166951668777. <http://doi.org/10.1177/2041669516687770>
- Haryu, E., & Kajikawa, S. (2012). Are higher-frequency sounds brighter in color and smaller in size? Auditory-visual correspondences in 10-month-old infants. *Infant Behavior and Development*, 35(4), 727–732.
<http://doi.org/10.1016/j.infbeh.2012.07.015>
- Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, 58(4), 1021–1034.
- Helm, C. E. (1964). Multidimensional Ratio Scaling Analysis of Perceived Color Relations. *Journal of the Optical Society of America*, 54(1 I), 256–262.
<http://doi.org/10.1364/JOSA.54.000256>
- Helmholtz, H. von. (1896). *Handbuch der Physiologischen Optik* (2nd ed). Hamburg: Voss.
- Hering, E. (1874). *Outlines of a Theory of the Light Sense*. Cambridge, Massachusetts: Harvard University Press.
- Hernández-Andrés, J., Romero, J., Nieves, J. L., & Lee, R. L. (2001). Color and spectral analysis of daylight in southern Europe. *Journal of the Optical Society of America A*, 18(6), 1325. <http://doi.org/10.1364/JOSAA.18.001325>
- Heywood, C., Gadotti, A., & Cowey, A. (1992). Cortical area V4 and its role in the

References

- perception of color. *The Journal of Neuroscience*, 12(10), 4056–4065.
<http://doi.org/10.1523/JNEUROSCI.12-10-04056.1992>
- Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint Measurement of Gloss and Surface Texture. *Psychological Science*, 19(2), 196–204.
<http://doi.org/10.1111/j.1467-9280.2008.02067.x>
- Honjyo, K., & Nonaka, M. (1970). Perception of White in a 10° Field. *Journal of the Optical Society of America*, 60(12), 1690. <http://doi.org/10.1364/JOSA.60.001690>
- Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, 13(4), 341–348.
<http://doi.org/10.1002/icd.364>
- Hunt, R. W. G., & Pointer, M. R. (2011). *Measuring Colour* (4th ed.). Wiley.
- Hurlbert, A. C. (1998). Computational models of color constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual Constancy: Why Things Look as They Do* (pp. 283–322). Cambridge University Press.
- Hurlbert, A. C. (2007). Colour constancy. *Current Biology*, 17(21), R906–R907.
<http://doi.org/10.1016/j.cub.2007.08.022>
- Hurlbert, A. C., & Ling, Y. (2005). If it's a banana, it must be yellow: The role of memory colors in color constancy. *Journal of Vision*, 5(8), 787–787.
<http://doi.org/10.1167/5.8.787>
- Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, 17(16), R623–R625.
<http://doi.org/10.1016/j.cub.2007.06.022>

References

- Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, 64(6, Pt.1), 384–404. <http://doi.org/10.1037/h0041403>
- Indow, T. (1988). Multidimensional studies of Munsell color solid. *Psychological Review*, 95(4), 456–479. <http://doi.org/10.1037/0033-295X.95.4.456>
- Indow, T., & Kanazawa, K. (1960). Multidimensional mapping of Munsell colors varying in hue, chroma, and value. *Journal of Experimental Psychology*, 59(5), 330–336. <http://doi.org/10.1037/h0044796>
- Ishihara, S. (2010). *Ishihara's test chart for colour deficiency*. Tokyo: Kanehara Trading INC.
- Jameson, D., & Hurvich, L. M. (1955). Some Quantitative Aspects of an Opponent-Colors Theory I: Chromatic Responses and Spectral Saturation. *Journal of the Optical Society of America*, 45(7), 546. <http://doi.org/10.1364/JOSA.45.000546>
- Kaldy, Z., & Blaser, E. (2009). How to Compare Apples and Oranges: Infants' Object Identification Tested With Equally Salient Shape, Luminance and Color Changes. *Infancy : The Official Journal of the International Society on Infant Studies*, 14(2), 222–243. <http://doi.org/10.1080/15250000802707088>
- Kaldy, Z., Blaser, E. A., & Leslie, A. M. (2006). A new method for calibrating perceptual salience across dimensions in infants: The case of color vs. luminance. *Developmental Science*, 9(5), 482–489. <http://doi.org/10.1111/j.1467-7687.2006.00515.x>
- Káldy, Z., & Kovács, I. (2003). Visual Context Integration is Not Fully Developed in 4-Year-Old Children. *Perception*, 32(6), 657–666. <http://doi.org/10.1068/p3473>

References

- Kaldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: integrating “what” and “where” information. *Developmental Science*, 6(3), 360–373. <http://doi.org/10.1111/1467-7687.00290>
- Kemler, D. G. (1983). Exploring and reexploring issues of integrality, perceptual sensitivity, and dimensional salience. *Journal of Experimental Child Psychology*, 36(3), 365–379. [http://doi.org/10.1016/0022-0965\(83\)90040-1](http://doi.org/10.1016/0022-0965(83)90040-1)
- Kemler, D. G., & Smith, L. B. (1978). Is there a developmental trend from integrality to separability in perception? *Journal of Experimental Child Psychology*, 26(3), 498–507. [http://doi.org/10.1016/0022-0965\(78\)90128-5](http://doi.org/10.1016/0022-0965(78)90128-5)
- Kemler, D. G., & Smith, L. B. (1979). Accessing similarity and dimensional relations: Effects of integrality and separability on the discovery of complex concepts. *Journal of Experimental Psychology: General*, 108(2), 133–150. <http://doi.org/10.1037/0096-3445.108.2.133>
- Kingdom, F. A. A., Bell, J., Haddad, C., & Bartsch, A. (2015). Perceptual scales for chromatic and luminance blur in noise textures. *Journal of Vision*, 15(9), 1–10. <http://doi.org/10.1167/15.9.6>
- Kleiner, M., Brainard, D. H., & Peilli, D. (2007). What’s new in Psychtoolbox-3? In *Perception ECVF* (p. 36).
- Knoblauch, K., Bieber, M. L., & Werner, J. S. (1998). M- and L-cones in early infancy: I. VEP responses to receptor-isolating stimuli at 4- and 8-weeks of age. *Vision Research*, 38(12), 1753–1764. [http://doi.org/10.1016/S0042-6989\(97\)00383-0](http://doi.org/10.1016/S0042-6989(97)00383-0)
- Knoblauch, K., & Maloney, L. (2009). Maximum Likelihood Conjoint Measurement.

References

- Modeling Psychophysical Data in R*, 1–21. <http://doi.org/10.1007/978-1-4614-4475-6>
- Knoblauch, K., & Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software*, 25(2), 1--26. <http://doi.org/10.1.1.204.8835>
- Knoblauch, K., & Maloney, L. T. (2012). *Modeling Psychophysical Data in R*. New York, NY: Springer New York. <http://doi.org/10.1007/978-1-4614-4475-6>
- Knoblauch, K., & Maloney, L. T. (2014). MLCM: Maximum Likelihood Conjoint Measurement. R package version 0.4.1.
- Knoblauch, K., Vital-Durand, F., & Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research*, 41(1), 23–36. [http://doi.org/10.1016/S0042-6989\(00\)00205-4](http://doi.org/10.1016/S0042-6989(00)00205-4)
- Komarova, N. L., & Jameson, K. A. (2013). A Quantitative Theory of Human Color Choices. *PLoS ONE*, 8(2), e55986. <http://doi.org/10.1371/journal.pone.0055986>
- Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, 94(4), 301–21. <http://doi.org/10.1016/j.jecp.2005.12.001>
- Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences of the United States of America*, 96(1), 307–12.
- Krantz, D. (1972). Visual Scaling. In D. Jameson & L. M. Hurvich (Eds.), *Visual Psychophysics*. Berlin: Springer-Verlag.

References

- Krantz, D. (1975). Color measurement and color theory: I. Representation theorem for Grassmann structures. *Journal of Mathematical Psychology*.
- Krantz, D., Luce, R., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement (Additive and Polynomial Representations)*, vol. 1. New York: Academic Press.
- Krantz, D., & Tversky, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 78(2), 151–169.
- Krauskopf, J., & Gegenfurtner, K. R. (1992). Color discrimination and adaptation. *Vision Research*, 32(11), 2165–2175. [http://doi.org/10.1016/0042-6989\(92\)90077-V](http://doi.org/10.1016/0042-6989(92)90077-V)
- Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, 22(9), 1123–1131. [http://doi.org/10.1016/0042-6989\(82\)90077-3](http://doi.org/10.1016/0042-6989(82)90077-3)
- Kusunoki, M., Moutoussis, K., & Zeki, S. (2006). Effect of Background Colors on the Tuning of Color-Selective Cells in Monkey Area V4. *Journal of Neurophysiology*, 95(5), 3047–3059. <http://doi.org/10.1152/jn.00597.2005>
- Laeng, B., Brennen, T., Elden, Å., Gaare Paulsen, H., Banerjee, A., & Lipton, R. (2007). Latitude-of-birth and season-of-birth effects on human color vision in the Arctic. *Vision Research*, 47(12), 1595–1607. <http://doi.org/10.1016/J.VISRES.2007.03.011>
- Land, E. H. (1977). The Retinex Theory of Color Vision. *Scientific American*, 237(6), 108–128. <http://doi.org/10.1038/scientificamerican1277-108>
- Land, E. H., & McCann, J. J. (1971). Lightness and Retinex Theory. *Journal of the*

References

- Optical Society of America*, 61(1), 1. <http://doi.org/10.1364/JOSA.61.000001>
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299–321.
[http://doi.org/10.1016/0885-2014\(88\)90014-7](http://doi.org/10.1016/0885-2014(88)90014-7)
- Lanthony, P. (1998). *Album Tritan* (2nd ed.). Paris: Laboratoire de la Vision des Couleurs.
- Laparra, V., Jiménez, S., Camps-Valls, G., & Malo, J. (2012). Nonlinearities and Adaptation of Color Vision from Sequential Principal Curves Analysis. *Neural Computation*, 24(10), 2751–2788. http://doi.org/10.1162/NECO_a_00342
- Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift Fur Naturforschung*, (36), 910–912.
- Laughlin, S. (1983). Matching coding to scenes to enhance efficiency. In O. J. Braddick & A. C. Sleight (Eds.), *Physical and biological processing of images* (pp. 42–52). Berlin, Heidelberg: Springer.
- Lee, B. B., Pokorny, J., Martin, P. R., Valberg, A., & Smith, V. C. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A*, 7(12), 2223.
<http://doi.org/10.1364/JOSAA.7.002223>
- Lee, H. C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America. A, Optics and Image Science*, 3(10), 1694–9. <http://doi.org/10.1364/JOSAA.3.001694>
- Lee, R. J., & Smithson, H. E. (2016). Low levels of specular support operational

References

- color constancy, particularly when surface and illumination geometry can be inferred. *Journal of the Optical Society of America A*, 33(3), A306.
<http://doi.org/10.1364/JOSAA.33.00A306>
- Li, C. J., Luo, M. R., & Hunt, R. W. G. (2000). A revision of the CIECAM97s model. *Color Research & Application*, 25(4), 260–266. [http://doi.org/10.1002/1520-6378\(200008\)25:4<260::AID-COL6>3.0.CO;2-9](http://doi.org/10.1002/1520-6378(200008)25:4<260::AID-COL6>3.0.CO;2-9)
- Lindsey, D. T., Brown, A. M., Reijnen, E., Rich, A. N., Kuzmova, Y. I., & Wolfe, J. M. (2010). Color channels, not color appearance or color categories, guide visual search for desaturated color targets. *Psychological Science*, 21(9), 1208–14.
<http://doi.org/10.1177/0956797610379861>
- Lockhead, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, 79(5), 410–419. <http://doi.org/10.1037/h0033129>
- Logvinenko, A. D. (2009). An object-color space. *Journal of Vision*, 9(11), 5–5.
<http://doi.org/10.1167/9.11.5>
- Logvinenko, A. D., & Tokunaga, R. (2011). Colour Constancy as Measured by Least Dissimilar Matching. *Seeing and Perceiving*, 24, 407–452.
<http://doi.org/10.1163/187847511X588746>
- Luce, R., & Tukey, J. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 32, 466–473.
- Maclaury, R. E., Hewes, G. W., Kinnear, P. R., Deregowski, J. B., Merrifield, R., Saunders, B. C., ... Roger, W. (1992). From Brightness to Hue An Explanatory Evolution. *Current Anthropology*, 33(2), 137–186.

References

- MacLeod, D. I. A. (2003). Colour Discrimination, Colour Constancy, and Natural Scene Statistics (The Verriest Lecture). In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and Defective Colour Vision*.
- MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69(8), 1183. <http://doi.org/10.1364/JOSA.69.001183>
- Macleod, D. I. A., & von der Twer, T. (2001). The Pleistochrome: Optimal Opponent Codes For Natural Colours. In R. Mausfeld & D. Heyer (Eds.), *Colour Perception: from Light to Object* (pp. 155–185). Oxford: Oxford University Press.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd editio). Lawrence Erlbaum Associates.
- Maloney, L. T., & Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision*, 3(8), 5. <http://doi.org/10.1167/3.8.5>
- MATLAB. (2016). *R2016a*. Natick, Massachusetts: The MathWorks Inc.
- McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. London: Chapman and Hall.
- McDermott, K. C., & Webster, M. A. (2012). Uniform color spaces and natural image statistics. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 29(2), A182-7. <http://doi.org/10.1364/JOSAA.29.00A182>
- Mollon, J. D. (1989). "Tho'' she kneel'd in that place where they grew..." The uses and origins of primate colour vision." *The Journal of Experimental Biology*, 146, 21–38.

References

- Mollon, J. D. (1999). Color vision: opsins and options. *Proceedings of the National Academy of Sciences of the United States of America*, 96(9), 4743–5.
<http://doi.org/10.1073/PNAS.96.9.4743>
- Mollon, J. D. (2003). The Origins of Modern Color Science. *The Science of Color*, 1–36.
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*. <http://doi.org/10.1016/j.visres.2017.11.001>
- Mollon, J. D., & Cavonius, C. (1987). The chromatic antagonisms of Opponent Process Theory are not the same as those revealed in studies of detection and discrimination. In G. Verriest (Ed.), *Colour Vision Deficiencies VIII*, (pp. 473–483). Junk Publishers.
- Moroney, N., Fairchild, M. D., Hunt, R. W. G., Li, C., Luo, M. R., & Newman, T. (2002). The CIECAM02 Color Appearance Model. In *IS&T/SID Tenth Color Imaging Conference*. Scottsdale, Arizona: Society for Imaging Science and Technology.
- Morrone, M., Burr, D., & Fiorentini, A. (1990). Development of Contrast Sensitivity and Acuity of the Infant Colour System. *Proceedings of the Royal Society B: Biological Sciences*, 242(1304), 134–139. <http://doi.org/10.1098/rspb.1990.0116>
- Morrone, M., Burr, D., & Fiorentini, A. (1993). Development of infant contrast sensitivity to chromatic stimuli. *Vision Research*, 33(17), 2535–2552.
[http://doi.org/10.1016/0042-6989\(93\)90133-H](http://doi.org/10.1016/0042-6989(93)90133-H)

References

- Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: the generalized linear mixed model. *Journal of Vision*, 12(11), 26. <http://doi.org/10.1167/12.11.26>
- Munsell, A. (1912). A pigment color system and notation. *The American Journal of Psychology*, 23(2), 236–244. <http://doi.org/10.2307/1412843>
- Nathans, J., Thomas, D., & Hogness, D. (1986). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science*, 232(4747), 193–202. <http://doi.org/10.1126/science.2937147>
- Nayatani, Y. (1998). A colorimetric explanation of the Helmholtz-Kohlrausch effect. *Color Research & Application*, 23(6), 374–378. [http://doi.org/10.1002/\(SICI\)1520-6378\(199812\)23:6<374::AID-COL5>3.0.CO;2-W](http://doi.org/10.1002/(SICI)1520-6378(199812)23:6<374::AID-COL5>3.0.CO;2-W)
- Newhall, S. M., Nickerson, D., & Judd, D. B. (1943). Final Report of the OSA Subcommittee on the Spacing of the Munsell Colors. *Journal of the Optical Society of America*, 33(7), 385. <http://doi.org/10.1364/JOSA.33.000385>
- Nichiporuk, N., Knoblauch, K., Abbatecola, C., & Shevell, S. (2017). The Lightness Distortion Effect: Additive Conjoint Measurement Shows Race Has a Larger Influence on Perceived Lightness of Upright than Inverted Faces. *Journal of Vision*, 17(10), 245. <http://doi.org/10.1167/17.10.245>
- Olkkonen, M., & Allred, S. R. (2014). Short-term memory affects color perception in context. *PloS One*, 9(1), e86488. <http://doi.org/10.1371/journal.pone.0086488>
- Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *Journal of Vision*, 9(12), 6–6. <http://doi.org/10.1167/9.12.6>

References

- Olkkonen, M., Witzel, C., Hansen, T., & Gegenfurtner, K. R. (2010). Categorical color constancy for real surfaces. *Journal of Vision*, 10(9), 16–16.
<http://doi.org/10.1167/10.9.16>
- Osorio, D., & Bossomaier, T. R. J. (1992). Human cone-pigment spectral sensitivities and the reflectances of natural surfaces. *Biological Cybernetics*, 67(3), 217–222.
<http://doi.org/10.1007/BF00204394>
- Osorio, D., & Vorobyev, M. (1996). Colour Vision as an Adaptation to Frugivory in Primates. *Proceedings of the Royal Society B: Biological Sciences*, 263(1370), 593–599. <http://doi.org/10.1098/rspb.1996.0089>
- Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), 8877–82. <http://doi.org/10.1073/pnas.0906172107>
- Pearce, B. (2015). *The asymmetries of colour constancy as determined through illumination discrimination using tuneable LED light sources*. Newcastle University.
- 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(2), e87989. <http://doi.org/10.1371/journal.pone.0087989>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
<http://doi.org/10.1163/156856897X00366>

References

- Pereverzeva, M., & Teller, D. Y. (2009). Simultaneous Color Contrast in 4-Month-Old Infants. *Perception*, 38(1), 30–43. <http://doi.org/10.1068/p6098>
- Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2013). Robust Correlation Analyses: False Positive and Power Validation Using a New Open Source Matlab Toolbox. *Frontiers in Psychology*, 3(January), 1–18. <http://doi.org/10.3389/fpsyg.2012.00606>
- Philipona, D. L., & O'Regan, J. K. (2006). Color naming, unique hues, and hue cancellation predicted from singularities in reflection properties. *Visual Neuroscience*, 23(3–4), 331–339. <http://doi.org/10.1017/S0952523806233182>
- Pieters, J. M. (1979). A conjoint measurement approach to color harmony. *Perception & Psychophysics*, 26(4), 281–286. <http://doi.org/10.3758/BF03199881>
- Pitchford, N. J., & Mullen, K. T. (2002). Is the Acquisition of Basic-Colour Terms in Young Children Constrained? *Perception*, 31(11), 1349–1370. <http://doi.org/10.1068/p3405>
- Poynton, C. (2012). *Digital video and HD: Algorithms and Interfaces*. Elsevier.
- Qi, L., Chantler, M. J., Siebert, J. P., & Dong, J. (2015). The joint effect of mesoscale and microscale roughness on perceived gloss. *Vision Research*, 115, 209–217. <http://doi.org/10.1016/j.visres.2015.04.014>
- R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Radonjic, A., Aston, S., Krieger, A., Cottaris, N. P., Brainard, D. H., & Hurlbert, A. C. (2016). Illumination discrimination in real and simulated scenes', 16, 1–18.

References

<http://doi.org/10.1167/16.11.2>

Radonjić, A., & Brainard, D. H. (2016). The nature of instructional effects in color

constancy. *Journal of Experimental Psychology: Human Perception and*

Performance, 42(6), 847–865. <http://doi.org/10.1037/xhp0000184>

Radonjic, A., Pearce, B., Aston, S., Krieger, A., Dubin, H., Cottaris, N. P., ... Hurlbert,

A. C. (2016). Illumination discrimination in real and simulated scenes. *Journal of*

Vision, 16(11), 2. <http://doi.org/10.1167/16.11.2>

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of

feature-label-order and their implications for symbolic learning. *Cognitive Science*,

34(6), 909–957. <http://doi.org/10.1111/j.1551-6709.2009.01092.x>

Redding, G. M., & Wallace, B. (1988). Components of prism adaptation in terminal and

concurrent exposure: Organization of the eye-hand coordination loop. *Perception*

& Psychophysics, 44(1), 59–68.

Regan, B. C., Julliot, C., Simmen, B., Viénot, F., Charles-Dominique, P., & Mollon, J.

D. (2001). Fruits, foliage and the evolution of primate colour vision. *Philosophical*

Transactions of the Royal Society of London. Series B, Biological Sciences,

356(1407), 229–83. <http://doi.org/10.1098/rstb.2000.0773>

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all.

Proceedings of the National Academy of Sciences, 102(23), 8386–8391.

<http://doi.org/10.1073/pnas.0503281102>

Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in

language. *Cross-Cultural Research*.

References

- Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology*.
- Robertson, A. R. (1990). Historical development of CIE recommended color difference equations. *Color Research & Application*, 15(3), 167–170.
<http://doi.org/10.1002/col.5080150308>
- Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., ... Vanduffel, W. (2012). Toward a Unified Theory of Visual Area V4. *Neuron*, 74(1), 12–29. <http://doi.org/10.1016/J.NEURON.2012.03.011>
- Rogers, M., Knoblauch, K., & Franklin, A. (2016). Maximum likelihood conjoint measurement of lightness and chroma. *Journal of the Optical Society of America A*, 33(3), A184. <http://doi.org/10.1364/JOSAA.33.00A184>
- Romero, J., García, J. A., del Barco, L. J., & Hita, E. (1993). Evaluation of color-discrimination ellipsoids in two-color spaces. *Journal of the Optical Society of America A*, 10(5), 827. <http://doi.org/10.1364/JOSAA.10.000827>
- Saarela, T. P., & Landy, M. S. (2015). Integration Trumps Selection in Object Recognition. *Current Biology*, 25(7), 920–927.
<http://doi.org/10.1016/j.cub.2015.01.068>
- Scarfe, P., & Glennerster, A. (2015). Using high-fidelity virtual reality to study perception in freely moving observers. *Journal of Vision*, 15(9), 3.
<http://doi.org/10.1167/15.9.3>
- Schein, S. J., & Desimone, R. (1990). Spectral properties of V4 neurons in the macaque.

References

- The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 10(10), 3369–89.
- Shapley, R., & Hawken, M. J. (2011). Color in the Cortex: single- and double-opponent cells. *Vision Research*, 51(7), 701–717.
<http://doi.org/10.1016/J.VISRES.2011.02.012>
- Shepard, R. N. (1995). The perceptual organization of colors: an adaptation to regularities of the terrestrial world. In J. H. Barlow, L. Cosmides, & J. Tooby (Eds.), *Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 495–532). Oxford: Oxford University Press.
- Shipp, S., & Zeki, S. (1985). Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. *Nature*, 315(6017), 322–5.
- Shipp, S., & Zeki, S. (2002). The functional organization of area V2, I: specialization across stripes and layers. *Visual Neuroscience*, 19(2), 187–210.
- Skelton, A., Catchpole, G., Abbott, J., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences*, (May 8, 2017).
<http://doi.org/10.1073/pnas.1612881114>
- Smith, L. B. (1979). Perceptual Development and Category Generalization. *Child Development*, 50(3), 705. <http://doi.org/10.2307/1128936>
- Smith, L. B. (1980). Development and the continuum of dimensional separability. *Perception & Psychophysics*, 28(2), 164–172. <http://doi.org/10.3758/BF03204343>
- Smith, L. B. (1983). Development of classification: The use of similarity and dimensional relations. *Journal of Experimental Child Psychology*, 36(1), 150–178.

References

[http://doi.org/10.1016/0022-0965\(83\)90064-4](http://doi.org/10.1016/0022-0965(83)90064-4)

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, 10(4), 502–532. [http://doi.org/10.1016/0010-0285\(78\)90009-9](http://doi.org/10.1016/0010-0285(78)90009-9)

Smith, T., & Guild, J. (1932). The C.I.E. colorimetric standards and their use. *Transactions of the Optical Society*, 33(3), 73–134. <http://doi.org/10.1088/1475-4878/33/3/301>

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15(2), 161–171. [http://doi.org/10.1016/0042-6989\(75\)90203-5](http://doi.org/10.1016/0042-6989(75)90203-5)

Smith, V. C., Pokorny, J., Davis, M., & Yeh, T. (1995). Mechanisms subserving temporal modulation sensitivity in silent-cone substitution. *Journal of the Optical Society of America A*, 12(2), 241. <http://doi.org/10.1364/JOSAA.12.000241>

Smithson, H. E. (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/10.1098/rstb.2005.1633>

Smithson, H. E. (2015). Perceptual organization of colour. In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization* (pp. 1–30). Oxford University Press.

Soja, N. N. (1994). Young Children's Concept of Color and Its Relation to the Acquisition of Color Words. *Child Development*, 65(3), 918–937.

Soto, F. A., Vucovich, L., Musgrave, R., & Ashby, F. G. (2015). General recognition theory with individual differences: a new method for examining perceptual and

References

- decisional interactions with an application to face perception. *Psychonomic Bulletin & Review*, 22(1), 88–111. <http://doi.org/10.3758/s13423-014-0661-y>
- Spehar, B., & Zaidi, Q. (2000). The crispening effect in luminance and colour. *Perception*, 29(ECVP Abstract Supplement).
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137–149. <http://doi.org/10.3758/BF03207704>
- Stockman, A., MacLeod, D. I. A., & Johnson, N. E. (1993). Spectral sensitivities of the human cones. *Journal of the Optical Society of America A*, 10(12), 2491. <http://doi.org/10.1364/JOSAA.10.002491>
- Sturges, J., & Whitfield, T. W. A. (1995). Locating basic colours in the munsell space. *Color Research & Application*, 20(6), 364–376. <http://doi.org/10.1002/col.5080200605>
- Sugita, Y. (2004). Experience in Early Infancy Is Indispensable for Color Perception. *Current Biology*, 14, 1267–1271. <http://doi.org/10.1016/j>
- Surridge, A. K., Osorio, D., & Mundy, N. I. (2003). Evolution and selection of trichromatic vision in primates. *Trends in Ecology & Evolution*, 18(4), 198–205. [http://doi.org/10.1016/S0169-5347\(03\)00012-0](http://doi.org/10.1016/S0169-5347(03)00012-0)
- Suttle, C. M., Banks, M. S., & Graf, E. W. (2002). FPL and sweep VEP to tritan stimuli in young human infants. *Vision Research*, 42(26), 2879–91.
- Takasaki, H. (1966). Lightness Change of Grays Induced by Change in Reflectance of Gray Background. *Journal of the Optical Society of America*, 56(4), 504.

References

<http://doi.org/10.1364/JOSA.56.000504>

Takasaki, H. (1967). Chromatic Changes Induced by Changes in Chromaticity of Background of Constant Lightness. *Journal of the Optical Society of America*, 57(1), 93. <http://doi.org/10.1364/JOSA.57.000093>

Taylor, C., Clifford, A., & Franklin, A. (2013). Color preferences are not universal. *Journal of Experimental Psychology: General*, 142(4), 1015–1027. <http://doi.org/10.1037/a0030273>

Teller, D. Y. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development*, 2, 135–153. [http://doi.org/10.1016/S0163-6383\(79\)80016-8](http://doi.org/10.1016/S0163-6383(79)80016-8)

Turner, R. S. (1993). Vision studies in Germany: Helmholtz versus Hering. *Osiris*, 8, 80–103.

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E.-J. (2016). Bayesian Inference for Kendall's Rank Correlation Coefficient. *The American Statistician*, 0–0. <http://doi.org/10.1080/00031305.2016.1264998>

Vazquez-Corral, J., Vanrell, M., Baldrich, R., & Tous, F. (2012). Color Constancy by Category Correlation. *IEEE Transactions on Image Processing*, 21(4), 1997–2007. <http://doi.org/10.1109/TIP.2011.2171353>

Volbrecht, V. J., & Werner, J. S. (1987). Isolation of short-wavelength-sensitive cone photoreceptors in 4-6-week-old human infants. *Vision Research*, 27(3), 469–478. [http://doi.org/10.1016/0042-6989\(87\)90094-0](http://doi.org/10.1016/0042-6989(87)90094-0)

von Helmholtz, H., & Southall, J. P. C. (1925). Helmholtz's Treatise on Physiological

References

- Optics. Translated from the Third German edition. *JAMA: The Journal of the American Medical Association*, 85(6), 461.
<http://doi.org/10.1001/jama.1925.02670060063036>
- von Kries, J. (1902). Chromatic adaptation. In D. L. MacAdam (Ed.), *Sources of color science* (pp. 109–119). Cambridge, MA: MIT.
- Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: color word learning as a gradual inductive process. *Cognition*, 127(3), 307–17.
<http://doi.org/10.1016/j.cognition.2013.01.010>
- Walraven, P. L. (1961). On the Bezold-Brücke Phenomenon. *Journal of the Optical Society of America*, 51(10), 1113. <http://doi.org/10.1364/JOSA.51.001113>
- Walsh, V., Carden, D., Butler, S. R., & Kulikowski, J. J. (1993). The effects of V4 lesions on the visual abilities of macaques: hue discrimination and colour constancy. *Behavioural Brain Research*, 53(1–2), 51–62.
- Walsh, V., & Kulikowski, J. (Eds.). (1998). *Perceptual constancy: Why things look as they do*. Cambridge University Press.
- Ward, G. J. (1994). The RADIANCE lighting simulation and rendering system. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (pp. 459–472). ACM.
- Ward, G. J., & Shakespeare, R. A. (1998). *Rendering with Radiance: The Art and Science of Lighting Visualization*. Morgan Kaufmann Publishers.
- Wass, S. V, Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults.

References

- Behavior Research Methods*, 45(1), 229–250. <http://doi.org/10.3758/s13428-012-0245-6>
- Webster, M. A. (1996). Human colour perception and its adaptation. *Network: Computation in Neural Systems*, 7(4), 587–634. <http://doi.org/10.1088/0954-898X/7/4/002>
- Webster, M. A., & Leonard, D. (2008). Adaptation and perceptual norms in color vision. *Journal of the Optical Society of America A*, 25(11), 2817. <http://doi.org/10.1364/JOSAA.25.002817>
- Webster, M. A., & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Research*, 37(23), 3283–3298. [http://doi.org/10.1016/S0042-6989\(97\)00125-9](http://doi.org/10.1016/S0042-6989(97)00125-9)
- Weiss, D., & Gegenfurtner, K. R. (2016). A comparison between illuminant discrimination and chromatic detection. In *39th European Conference on Visual Perception (ECVP)* (Vol. 45, p. 214). <http://doi.org/10.1177/0301006616671273>
- Werner, J. S., & Scheffrin, B. E. (1993). Loci of achromatic points throughout the life span. *Journal of the Optical Society of America A*, 10(7), 1509. <http://doi.org/10.1364/JOSAA.10.001509>
- Werner, J. S., & Walraven, J. (1982). Effect of chromatic adaptation on the achromatic locus: The role of contrast, luminance and background color. *Vision Research*, 22(8), 929–943. [http://doi.org/10.1016/0042-6989\(82\)90029-3](http://doi.org/10.1016/0042-6989(82)90029-3)
- Whittle, P. (1992). Brightness, discriminability and the “Crispening Effect.” *Vision Research*, 32(8), 1493–1507. [http://doi.org/10.1016/0042-6989\(92\)90205-W](http://doi.org/10.1016/0042-6989(92)90205-W)

References

- Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, 72(2), 125–166. [http://doi.org/10.1016/S0010-0277\(99\)00035-9](http://doi.org/10.1016/S0010-0277(99)00035-9)
- Winkler, A. D., Spillmann, L., Werner, J. S., & Webster, M. A. (2015). Asymmetries in blue–yellow color perception and in the color of ‘the dress.’ *Current Biology*, 25(13), R547–R548. <http://doi.org/10.1016/j.cub.2015.05.004>
- Witzel, C., Maule, J., & Franklin, A. (2013). Focal colors as perceptual anchors of color categories. *Journal of Vision*, VSS2013 abstracts.
- Witzel, C., Racey, C., & O'Regan, J. K. (2017). The most reasonable explanation of “the dress”: Implicit assumptions about illumination. *Journal of Vision*, 17(2), 1. <http://doi.org/10.1167/17.2.1>
- Witzel, C., Sanchez-Walker, E., & Franklin, A. (2013). The development of categorical colour constancy. *Perception*, 42(ECVP Abstract Supplement), 19.
- Witzel, C., van Alphen, C., Godau, C., & O'Regan, J. K. (2016). Uncertainty of sensory signal explains variation of color constancy. *Journal of Vision*, 16(15), 8. <http://doi.org/10.1167/16.15.8>
- Wood, S. N. (2015). *Core Statistics*. Cambridge, UK: Cambridge University Press.
- Woods, R. J., & Wilcox, T. (2010). Covariation of color and luminance facilitate object individuation in infancy. *Developmental Psychology*, 46(3), 681–690. <http://doi.org/10.1037/a0019161>
- Worthey, J. A. (1985). Limitations of color constancy. *Journal of the Optical Society of America A*, 2(7), 1014. <http://doi.org/10.1364/JOSAA.2.001014>

References

- Wright, W. D. (1929). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4), 141–164.
<http://doi.org/10.1088/1475-4878/30/4/301>
- Wyszecki, G., & Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae* (2nd ed.). Wiley.
- Xiao, Y., Wang, Y., & Felleman, D. J. (2003). A spatially organized representation of colour in macaque cortical area V2. *Nature*, 421(6922), 535–539.
<http://doi.org/10.1038/nature01372>
- Yang, J., Kanazawa, S., Yamaguchi, M. K., & Kuriki, I. (2013). Investigation of color constancy in 4.5-month-old infants under a strict control of luminance contrast for individual participants. *Journal of Experimental Child Psychology*, 115(1), 126–36. <http://doi.org/10.1016/j.jecp.2012.11.013>
- Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Vision Research*, 41(20), 2581–2600.
[http://doi.org/10.1016/S0042-6989\(01\)00143-2](http://doi.org/10.1016/S0042-6989(01)00143-2)
- Young, T. (1800). Outlines of experiments and inquiries respecting sound and light. *Philosophical Transactions of the Royal Society*, 90(106–150).
- Young, T. (1802). The Bakerian Lecture. On the Theory of Light and Colours. *Philosophical Transactions of the Royal Society of London*, 92(12–48).
- Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Research*, 26(6), 847–855.
[http://doi.org/10.1016/0042-6989\(86\)90143-4](http://doi.org/10.1016/0042-6989(86)90143-4)

References

Zchaluk, K., & Foster, D. H. (2009). Model-free estimation of the psychometric function. *Attention, Perception, & Psychophysics*, 71(6), 1414–1425.

<http://doi.org/10.3758/APP.71.6.1414>

Zemach, I., Chang, S., & Teller, D. Y. (2007). Infant color vision: Prediction of infants' spontaneous color preferences. *Vision Research*, 47(10), 1368–1381.

<http://doi.org/10.1016/j.visres.2006.09.024>