Development of visual perception and its implication for education and design

Teresa Tang

Thesis submitted for the degree of Doctor of Philosophy

University of Sussex

September 2021

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. The thesis conforms to an 'article format' in which the first chapter presents an introductory overview chapter of the relevant literature and an outline of the empirical work of the thesis. The remaining chapters consist of four papers written in a style appropriate for publication. One of these chapters has been published, and three are prepared for journal submission. The final chapter is a discussion of the overall contribution of the thesis to the field, future research and conclusions.

Chapters and author contributions

Chapter 1 provides an overview of the relevant literature, concepts and theories, and a summary of the papers.

Chapter 2 is written in American English and published in *Behavior Research Methods* as:

> Tang, T., Álvaro, L., Alvarez. J., Maule., J., Skelton, A., Franklin, A., & Bosten, J. (2021). *ColourSpot*, a novel gamified tablet-based test for accurate diagnosis of color vision deficiency in young children. *Behavior Research Methods,*1-13. https://doi.org/10.3758/s13428-021-01622-5

Author contributions: TT, LA, JA, AF and JB designed the research, TT collected the data, JM and AS did the iPad colorimetry calibration, TT, JA, and JB analysed the data, TT, LA, JB and AF wrote the paper.

Chapter 3 has been written in a style appropriate for *Journal of Educational Psychology*:

> Tang, T. & Franklin, A. The impact of colour vision deficiency in children and adolescents on self-reported wellbeing, educational engagement and ability to complete school tasks.

Author contributions: TT and AF designed the research, TT collected and analysed the data, TT and AF wrote the paper.

Chapter 4 has been written in a style appropriate for *Infancy*:

> Tang, T., Skelton, A., Bosten, J. & Franklin, A. The contribution of various pattern characteristics and image statistics to visual preference in infants and toddlers.

Author contributions: TT and AF designed the research, TT collected and analysed the data, AS and JB coded the image statistics toolbox, TT and AF wrote the paper.

Chapter 5 has been written in a style appropriate for *Perception:*

> Tang, T., Skelton, A., Bosten, J., Alvarez, J. & Franklin, A. Infants do not look longer at abstract images with natural chromatic scene statistics.

Author contributions: TT, AS, JB, JA designed the research, TT collected the data, TT and AS analysed the data, TT and AF wrote the paper.

Chapter 6 provides the overall contribution of the thesis, ideas for future research and conclusions.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Anna Franklin for her expertise, encouragement and kindness throughout the years. The thesis would never have reached this stage without her patience and persistence. Thank you for always being super supportive, super encouraging and sharing your optimism superpower. Thank you also to my secondary supervisor, Dr Jenny Bosten for her patience and insight. I am also grateful to all my colleagues at the Sussex Colour Group and Baby Lab for their company and support. Special thanks to Dr Alice Skelton, for always having faith in me and being my fellow go-to developmental psychologist to share all things developmental; and Dr-to-be Lucy Somers, for giving me motivational boosts and sharing her thesis writing music playlist. Thank you to my colleagues: Dr John Maule, Dr-to-be Simon Saryazdi, Dr-to-be Ian Pennock, Dr-to-be Alex Swartz, and everyone else I've worked with during these years for all the laughs, tea-and-biscuit-cake breaks and non-work-related office chats which have made my PhD journey more endurable. Although I never planned to do a PhD and this opportunity came very suddenly, knowing I would be well-supported in our lovely lab environment played a big decision for me to go ahead and I couldn't imagine doing it anywhere else.

I am also grateful to our collaborators, Dr Leticia Álvaro, who I hope to write another paper together with one day, as well as Dr James Alvarez, for his research software developing skills, as without him our experiments would not have happened.

Thank you to all the schools who participated in our research. Thank you to every parent, every child, and every baby I have met and tested during my time at the Sussex Baby Lab. Working with you little baby and child scientists was always the highlight of my day.

Thank you to my friends and family, who will never read this and/or do not understand what a PhD is, but I am always thankful for their presence and support regardless of what I do.

And lastly, I would like to thank and acknowledge myself. Thank you for persevering and persistently carrying on. You did it.

This PhD is internally funded by the School of Psychology, University of Sussex and externally funded by Cosatto Ltd.

COVID-19 impact statement

In March 2020, the first COVID-19 lockdown occurred in the UK. The University of Sussex was closed, all in-person testing was ceased, and everyone worked remotely. This prevented me from accessing equipment and data located in the lab, such as eye-tracking, the anomaloscope (the gold standard test for assessing colour vision) and calibration data. At this point in time, I was 17 months into my PhD and had completed testing for Papers 1, 3 and 4. Data collection for Paper 1 was complete, analysed and was in the process of writing a paper for publication. Data for Paper 4 were also collected and analysed. Data for Paper 3 were initially collected as an audit report for Cosatto Ltd to compare infants and toddlers looking when viewing their design patterns. It was not initially part of the thesis plan, which was to record images of infant's visual environment using infant head-mounted cameras and measure their image statistics. However, due to restrictions caused by COVID-19 this study was not possible, and the data collected for Cosatto's audit report was analysed and taken further by implementing an image statistic toolbox on design patterns to investigate how this may contribute to visual preference in infants and toddlers (Paper 3). Although this was not initially planned, Paper 3 has still revealed interesting findings that benefit and contribute to developmental science.

Prior to COVID-19, there were plans for in-person testing for Paper 2. The original plan was to test a large group of 11-16 year old children in secondary schools on the anomaloscope and compare it with *ColourSpot*, our new tablet-based colour vision test for children and in-person questionnaires about child wellbeing and quality of life. Prior to COVID-19, schools were contacted and arranged for school testing. However, these plans were re-adapted after COVID-19, where all testing tasks were moved online and only those with an eligible iPad could participate in order to download *ColourSpot*. Recruiting young people online was extremely difficult most likely due to the impact of COVID-19, i.e., school closures, online fatigue and COVID-related stress. For example, only 29% of participants who signed up fully completed the tasks. Hence, the impact of COVID-19 meant the scale of Paper 2 had to be minimised to be feasible. However, Paper 2 still makes an important contribution to science by identifying school tasks that children with colour vision deficiency find difficult, and by showing how *ColourSpot* enables the effect of colour vision deficiency on children to be tested remotely.

COVID-19 has significantly impacted my original research plans and prevented further research opportunities, but despite these circumstances I have been creative, adaptive and resilient to develop new methods of working, and have still provided papers that contribute to knowledge and science.

UNIVERSITY OF SUSSEX

TERESA TANG

PHD PSYCHOLOGY

DEVELOPMENT OF VISUAL PERCEPTION AND ITS IMPLICATIONS ON EDUCATION AND DESIGN

Summary

Developmental colour science has investigated how colour perception develops, but little is known about the impact of being colour vision deficient during development and the role of colour in developmental aesthetics. This thesis aims to contribute to theories of perceptual development and to apply research on the development of colour and visual perception to the fields of education and design. This thesis addresses two topics: colour vision deficiency (CVD) in children, and the role of colour and other visual properties in complex images on infants' and toddlers' visual preferences.

Paper 1 develops and validates *ColourSpot*, a new gamified psychophysical test of CVD that is suited for remote and self-administered diagnosis from 4 years old. The paper shows that *ColourSpot* diagnoses CVD better than other paediatric CVD tests and is a potential tool for improving diagnosis of CVD in children. Paper 2 uses *ColourSpot* to investigate the impact of CVD on education and wellbeing in 11–16 year olds, identifying educational tasks that children with CVD find challenging.

Paper 3 investigates the role of colour and other visual features in complex abstract images on infants' and toddlers' visual preferences. Infants show preferences for faces or odd-one-out features and the dominant colour in an image contributes to visual preference, but not other chromatic and spatial image statistics. Paper 4 further investigates the role of natural chromatic statistics in visual preference and finds that infants do not look longer at abstract images with natural chromatic distributions, contrary to the efficient encoding hypothesis.

The findings produce a gamified optometric test of CVD that can improve diagnosis of CVD in children and provides a better understanding of how to support children with CVD. The thesis also uses an image statistics approach to further understand how the visual system develops and tunes into the characteristics of the visual world and will have implications on guiding designers in producing designs optimised for infants and children.

Table of Contents

Chapter 1
Introduction and thesis overview

Colour vision can be defined as the ability to make discriminations based on the wavelength composition of the light independent of its intensity (De Valois & De Valois, 2000; Hilbert, 1992). Existing research has indicated that colour vision occurs via the comparison of signals sent by the short (S), medium (M) and long (L) wavelength sensitive cones in the retina which is then processed in the brain (Von Helmholtz, 1867; Young, 1802). Colour can also be seen as a visual perceptual characteristic defined in terms of hue (the dominant wavelength of colour), brightness (the amount of luminance and lightness in a colour) and saturation (the intensity of colour) (Burns & Shepp, 1988; Hunter & Pointer, 2011; Wyszecki & Stiles, 1982)).

Colour is an important feature of our visual experience that influences human mind and behaviour in a number of ways. Colour helps us recognise and search for objects, infer about their properties (e.g., the banana is ripe), communicate (e.g., pass me the red one), and informs us about our environment (e.g., the colour changes in daylight illumination tells us the time of day). Humans have an emotional connection to colour: they have aesthetic preferences for single colours (Hurlbert & Ling, 2007; Palmer & Schloss, 2010; Skelton & Franklin, 2020; C. Taylor et al., 2013), colour affects human appreciation of art (Albers et al., 2020; Altmann et al., 2021; Nakauchi et al., 2018; Nascimento et al., 2017, 2021), and colour has conceptual and abstract associations (Palmer & Schloss,

2010). Colour is also used as a signal (e.g., the red stop sign), and is an important factor in design and marketing (Aslam, 2006; Kauppinen-Räisänen & Luomala, 2010). It has also been argued that the presence of certain colours can make us more creative in tasks (Lichtenfeld et al., 2012; Studente et al., 2016), improve performance on IQ tests (Elliot et al., 2007; Larsson & von Stumm, 2015), or even affects the sexual desirability of a person wearing that colour (Elliot et al., 2010; Elliot & Niesta, 2008; Pazda et al., 2012). Given the importance of colour for mind and behaviour, it is critical as psychologists that we understand how humans see and respond to colour. Research on human colour vision and perception also has potential to give insight into theoretical questions such as how the brain processes visual information and role of experience in perception (Conway et al., 2010; Gegenfurtner, 2003; Hubel & Wiesel, 2004).

This thesis' main focus is on the development of colour perception. A long tradition of research on perceptual development has shown how understanding the development of a perceptual process can be useful for testing theories of perception (Fantz, 1961; Fantz & Miranda, 1975; Fantz & Nevis, 1967). Understanding how infants and children see and respond to colour is also important for understanding their response to their environment more broadly. Colour may be especially influential for infants and children given the highly colour coded nature of their world, with children exposed to highly coloured toys, books, television and educational materials. This thesis has the aim of

contributing to theories of perceptual development and applying research on the development of colour and visual perception, so that it is useful to the fields of education and design. The thesis presents four papers that address both theoretical and applied questions. Two of the papers investigate colour vision deficiency (‘colour blindness’) in children and aim to develop and evaluate an easily accessible and child-friendly test that diagnoses colour vision deficiency in children accurately (Paper 1); and investigates whether colour vision deficiency has an impact on children’s wellbeing and educational engagement (Paper 2). The other two papers investigate colour perception in infancy and early childhood, and its role in infant and toddler visual preference. Paper 3 investigates the contribution of colour and other visual features to infant and toddler pattern preference. Paper 4 asks whether infants prefer to look at abstract images with natural colour distributions. These latter two papers both aim to investigate theoretical questions on the role of scene statistics (the statistical regularities of visual features in natural images) in perceptual development, but also have the applied aim of identifying principles that could lead to optimal pattern design for infants and young children.

The thesis follows the papers style format and following this current overview chapter the thesis will present each paper in order. The current chapter will introduce and review the broader research context from which the thesis originates and explain the important concepts that the thesis draws on. The

chapter will then explain the rationale and importance of the questions of the thesis and will give a summary of each of the papers.

1.1 Introduction to the relevant literature, concepts and theories

The physiology of colour vision

Colour is an intrinsic property of our visual experience, and there has been a massive research effort into understanding the underlying physiology and neural basis of colour vision. This research traces back hundreds of years when scientists Young (1802) and Von Helmholtz (1867) proposed the trichromatic colour vision theory. This theory suggested that normal human colour vision is perceived through the combination of signals from the three cone photoreceptors: short (S), medium (M) and long (L) wavelength sensitive cones in the retina, which respond to short, medium, and long wavelengths of light (Von Helmholtz, 1867; Young, 1802). This theory was built on by Hering (1874) who proposed the opponent process theory which suggests colour is mediated by three opposing mechanisms: red-green, blue-yellow and black-white.

The early theories of colour vision have heavily influenced how we believe colour is processed in the brain today. Modern neuroscience has provided evidence for the existence of the three cone types, and it has also been found that the retinal ganglion cells contain three independent channels (known as the cardinal directions of colour space) which convey colour information from the

retina to the brain (Krauskopf et al., 1982). These include a luminance channel (L+M), where the L and M cone signals are added to compute the luminance intensity of a stimulus; a ‘red-green’ opponent channel (L-M), where the difference of the L and M cone signals are subtracted to compute a ‘red-green’ component of a stimulus; and a ‘blue-yellow’ opponent channel (S-(L+M)), where the L and M cone signals are subtracted from the S cone signal to compute a ‘blue-yellow’ component of a stimulus (De Valois et al., 1966; Gegenfurtner, 2003; Krauskopf et al., 1982). Note, the terms ‘red-green’ and ‘blue-yellow’ are commonly used to refer to the cardinal mechanisms, but do not refer to the colour appearance of the mechanisms which in fact correspond to cherry-teal and violet-chartreuse. The three cardinal mechanisms coincide with three distinct independent retino-geniculo-cortical pathways, which are pathways responsible for projecting information from the lateral geniculate nucleus (LGN) to the primary visual cortex (V1) (De Valois et al., 1966; Derrington et al., 1984; Krauskopf et al., 1982). The cells in the magnocellular layer of the LGN receive inputs from the L and M cones (L+M), and are most sensitive to luminance information; whereas cells in the parvocellular layers of the LGN are most sensitive to ‘red-green’ information and cells in the koniocellular layers of the LGN receive inputs from the S cones and are most sensitive to ‘blue-yellow’ information (Buchsbaum & Gottschalk, 1983; Derrington et al., 1984; Gegenfurtner, 2003; Gegenfurtner & Kiper, 2003; Kaplan et al., 1990; J. Lee & Stromeyer, 1989). Beyond the LGN, colour-selective cells have been found in the V1, V2, V3 and V4, as well as regions of

the inferior temporal cortex (Conway, 2009; Conway et al., 2010; Zeki & Bartels, 1999).

The physiological basis of colour vision and the cardinal mechanisms of colour have been used to derive colour spaces that quantify the visual system's physiological response to different colours. This thesis uses a physiological colour space called the MacLeod-Boynton chromaticity diagram (MacLeod & Boynton, 1979) which represents the visual system's response to the cone excitations of the L, M and S cones. The MacLeod-Boynton chromaticity diagram maintains the assumption that the S cone does not contribute to luminance (L+M). Hence luminance remains constant, and colour is represented in an isoluminant chromaticity plane. This occurs by scaling the cone fundamentals (spectral sensitivities) of the L, M and S cones into a chromaticity diagram where L+M sums to the luminosity function (V. Smith & Pokorny, 1975, 1996; Stockman et al., 1993). The luminosity function is a measure of the spectral sensitivity of brightness (Lennie et al., 1993). If plotted in cartesian coordinate space, the MacLeod-Boynton chromaticity diagram plots L/(L+M) on the horizontal x axis, representing the L and M cone excitations (red-green) and S/(L+M) on vertical y axis, representing the S cone excitation (blue-yellow). See Figure 1.1.

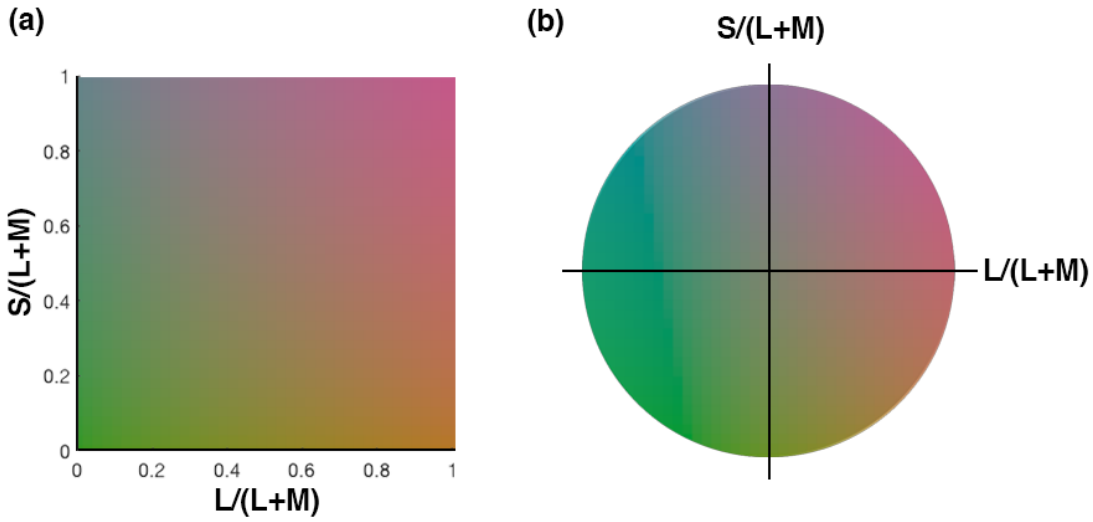


Figure 1.1. The MacLeod-Boynton chromaticity diagram represented in (a) cartesian space and (b) in polar space, showing the cardinal directions of colour, with the red-green axis (L/(L+M)) and the blue-yellow axis (S/(L+M). Please note that L/(L+M) can also be denoted as L-M and S/(L+M) can also be denoted as S-(L+M).

The thesis uses the MacLeod-Boynton colour space as it is a popular physiological colour space that is appropriate for the investigation of colour vision deficiency. The MacLeod-Boynton colour space has also been used in prior investigations of chromatic scene statistics, which the latter papers of this thesis also investigate. A detailed overview of different physiological and perceptual colour spaces can be found in other texts (Hunter & Pointer, 2011; Wyszecki & Stiles, 1982).

Colour vision deficiency

Although the above outline of trichromatic colour vision is the norm, some individuals are colour vision deficient. Colour vision deficiency (CVD) is the most common genetic visual abnormality (Birch, 2012), and involves abnormality in, or absence of, one of the L, M or S cone photoreceptor types. Protan, deutan and tritan refer to types of CVD relating to the L, M and S cone respectively. Tritan deficiency, which is an abnormality (tritanomaly) or absence of the S cone photopigment (tritanopia), is extremely rare with a prevalence of 0.0001% (Wright, 1952). The most common type of CVD is commonly referred as 'red-green' CVD which refers to an abnormality or absence of the L or M cone photopigments. Red-green CVD has a prevalence of approximately 8% in males and 0.4% in females in Europeans (Birch, 2012). This prevalence appears to vary worldwide amongst different ethnic populations with a lower prevalence of approximately 4% to 6.5% in east Asian males and approximately 3.6% to 4.8% in African males (Birch, 2012; Chia et al., 2008; Oduntan et al., 2019; Ugalahi, Fasina, Ogun, et al., 2016; Woldeamanuel & Geta, 2018). Red-green CVD is more common in males than females because it is an X-linked recessive inheritance disorder, meaning it's a genetic condition associated with variations on the X chromosome gene which encode the L and M pigment protein (Nathans et al., 1986; Neitz & Neitz, 1995, 2000). There are four sub-types of red-green CVD. Anomalous trichromacy is considered a milder form of red-green CVD and refers to an abnormality in the L or M cone photopigment's spectral sensitivity. There are two types of anomalous trichromacy: protanomaly, which refers to an abnormality in the L cone photopigment and

deuteranomaly, which refers to an abnormality in the M cone photopigment. The other type of red-green CVD is dichromacy, which is considered a severe form of red-green CVD and refers to an absence of either of the L or the M photopigment. There are also two types of dichromacy: protanopia, which refers to the absence of the L cone photopigment and deuteranopia, which refers to the absence of the M cone photopigment (Neitz & Neitz, 2000; Parry, 2015; Sharpe et al., 1999).

There has been a huge research effort to develop accurate diagnostic tests of CVD. These tests diagnose CVD using coloured stimuli that lie along dichromatic “confusion lines”. Dichromatic confusion lines are lines in colour space which dichromats will be unable to discriminate due to their missing respective cone type (Lakowski, 1969; Moreira et al., 2018). There are three types of confusion lines (protan, deutan and tritan), which represent the spectral sensitivity of the missing fundamental L, M or S cone of protanopes, deuteranopes and tritanopes respectively. Although confusion lines are based on dichromats, anomalous trichromats will also find stimuli on dichromatic confusions lines difficult to discriminate but this may vary across anomalous trichromat individuals as each will have a varied individual spectral sensitivity of the L or M cone (Bosten, 2019). This is one of the reasons why diagnosing CVD severity is a challenge for diagnostic tests. At present, the anomaloscope is regarded as the gold standard test of assessing CVD in adults (Fanlo Zarazaga et al., 2019). The anomaloscope is an optical instrument where individuals are

asked to match two coloured lights until they are perceived as equal. This involves mixing different levels of red and green monochromatic light to match a yellow monochromatic light. However, the anomaloscope is unlikely to be used as an instrument by clinical practitioners due to its high cost (B. Cole, 2007), and it is too demanding for young children (Verriest, 1982), having only been successfully used with children older than 7 years old (Jurasevska et al., 2014). Other tests, such as the Ishihara test (Ishihara, 1917), which is a book of pseudoisochromatic test plates, are more commonly used by practitioners (B. Cole, 2007; Fanlo Zarazaga et al., 2019). In addition to the development of tests, CVD research has investigated the nature of perceptual colour experience and colour naming in adults with CVD (Álvaro et al., 2015; Lillo, Moreira, et al., 2014; Tregillus et al., 2021), methods of simulating how CVD individuals see colour (Lillo, Álvaro, et al., 2014), the types of colour discriminations and tasks that are difficult for those with CVD (Álvaro et al., 2017; Steward & Cole, 1989), and the impact of CVD on adults' quality of life, career and wellbeing (Barry et al., 2017; Chan et al., 2014; B. Cole, 2004; Steward & Cole, 1989; Tagarelli et al., 2004).

There is much less work that considers CVD diagnosis, experience and impact in children. Several paediatric tests of CVD have been developed (reviewed in Table 2.1 of Paper 1 of the thesis). However, these tests have several limitations, particularly when used with young children. First, mass screening of CVD with the tests can be expensive as the tests require specialised equipment

and a trained administrator to administer the test. Second, for some tests there is little data on sensitivity and specificity when used on children (see Table 2.1), or the data indicate high rates of error (Tekavčič Pompe & Stirn Kranjc, 2012). Third, many paediatric CVD tests are not adapted to the capabilities of young children and require an understanding of numbers, shapes, or animals, which young children and children with additional and special educational needs will find challenging (Tekavčič Pompe, 2020; Tekavčič Pompe & Stirn Kranjc, 2012). Lastly, many of these tests do not have a control measure to differentiate between non-visual and colour vision errors, which is particularly important for young children who have limited attention. Hence, although there are a number of diagnostic CVD tests, there is still a need for an accurate, accessible and child-friendly paediatric CVD test which will address these problems. Such a test would enable better screening of CVD in children. Many countries, including the UK, do not have an early colour vision screening programme or routinely screen for CVD (Atowa et al., 2019; Azizoğlu et al., 2017; Ciner et al., 1999; B. Cole, 2015; Department of Health, 2009; Holroyd & Hall, 1997; Hopkins et al., 2013; Jadhav et al., 2017; Ramachandran et al., 2014; Stewart-Brown & Haslum, 1988; WHO Programme for the Prevention of Blindness and Deafness, 2003). This means that many children who have CVD are unaware of their condition, with 80% of children in the UK unaware of their CVD condition when they enter secondary school (Albany-Ward, 2011). The impact of CVD in children is also unclear. Some studies highlight that many school tasks and educational materials critically rely on normal colour vision (Mashige, 2019; Torrents et al., 2011; Ugalahi, Fasina, & Ogun, 2016), and suggest that CVD

increases the risk of emotional and behavioural difficulties in children (Thomas et al., 2018; Thuline, 1964). However, other studies argue against the negative impact of CVD in children (Cumberland et al., 2004; Nithiyaananthan et al., 2020; Ramachandran et al., 2014), which is at odds with the evidence that CVD significantly negatively impacts the quality of life in adults (Barry et al., 2017; Long et al., 2015; Steward & Cole, 1989; Stoianov et al., 2019). In sum, further research on CVD in children is needed, both to develop a diagnostic paediatric CVD test that could be used to screen for CVD in young children, and to clarify the impact of CVD on children's education, wellbeing and quality of life. Papers 1 and 2 of this thesis aim to address this need.

Colour perception, its development and the role of natural scenes

Although colour vision depends on cone-opponent mechanisms, these do not capture adults' perceptual experience of colour which is perceived along the perceptual dimensions of hue, saturation and brightness (Burns & Shepp, 1988; Hunter & Pointer, 2011; Wyszecki & Stiles, 1982). Adult colour discrimination sensitivity, which is plotted as a "discrimination ellipse" in MacLeod-Boynton colour space, also aligns not with the cardinal axes of colour vision, but with the blue-yellow daylight axis, which is an axis in colour space that connects colours that are bluish and orange-yellowish in appearance (Bosten et al., 2015; Panorgias et al., 2012). This blue-yellow daylight axis also closely corresponds to the daylight locus which is a line in colour space showing all the chromaticities of natural daylight illumination changes throughout the day

(Granzier & Valsecchi, 2014; Panorgias et al., 2012; Shepard, 1992). When the chromaticities of natural scene images are plotted in the MacLeod-Boynton colour space, they form a distribution that closely aligns with the blue-yellow daylight axis (Burton & Moorhead, 1987; McDermott & Webster, 2012; Shevell & Kingdom, 2008; Van Hateren & Ruderman, 1998; Webster et al., 2007; Webster & Mollon, 1997), and that also closely corresponds to the adult discrimination ellipse, where adults have poorer colour discrimination along the blue-yellow daylight axis (Álvaro et al., 2017; Bosten et al., 2015; Krauskopf & Gegenfurtner, 1992; Pearce et al., 2014). This finding suggests that colour perception may be calibrated to the statistics of the chromaticities and illumination of natural scenes, which fits with the broader theory that vision is optimally tuned and encodes the statistical regularities of the natural world (Olshausen & Field, 2000; Simoncelli & Olshausen, 2001; Webster et al., 2007).

There are a number of other perceptual colour processes in adults. For example, colour constancy is a perceptual process that enables the perceiver to keep the colour of objects constant despite changes in the object's illumination (Foster, 2011; Hurlbert, 2007). There has been an immense research effort to identify the cues that the brain uses in order for it to be able to keep colours perceptually constant in this way (e.g., see Smithson (2005) for a review). There has also been a large amount of research on the phenomenon of colour preference, whereby humans have preferences for single colour patches (McManus et al., 1981; Ou et al., 2004; Palmer et al., 2013; Palmer & Schloss,

2010). Decades of studies have shown that for adults, blue is commonly preferred and dark yellow-green is commonly disliked (Hurlbert & Ling, 2007; McManus et al., 1981; Palmer & Schloss, 2010; C. Taylor et al., 2013). Several theories have been proposed to account for this finding, from a model that accounts for colour preference in terms of cone-opponent colour vision (Hurlbert & Ling, 2007), to the ecological valence theory that accounts for colour preference in terms of valence associations of colours and objects (Palmer & Schloss, 2010). Related to the topic of colour preference, there has been investigation of adults' preference for the chromaticity of art. When adults are shown colour manipulated paintings, they will prefer the painting with the original colour distribution compared to manipulated colour versions, even when the paintings are scrambled (Albers et al., 2020; Altmann et al., 2021; Nascimento et al., 2017, 2021). One study has also found that adults will prefer abstract Mondrian images (overlapping rectangles of varying size, Land (1983)) containing chromaticities that fall along the natural blue-yellow daylight axis than any other axis in colour space (Juricevic et al., 2010). This again suggests that the chromaticities and illumination of natural scenes govern aspects of colour perception and influence not just colour discrimination sensitivity (Bosten et al., 2015; Krauskopf & Gegenfurtner, 1992; Pearce et al., 2014), but also govern aesthetic preference.

Research has also investigated the development of colour perception of infants and children. A body of rigorous psychophysical infant studies conducted over

two decades has shown that newborns have very limited colour vision (Adams, 1987; Bornstein, 1978; Hendrickson & Yuodelis, 1984; Morrone et al., 1990, 1993; Yuodelis & Hendrickson, 1986) and that trichromatic colour vision develops rapidly over the first few months of life, with the L-M cone opponent process developing around 4-6 weeks before the S-(L+M) cone opponent process (see Teller (1998) for a comprehensive review of infant colour vision in the first 4 months of life). Once trichromatic colour vision has developed, chromatic discrimination threshold sensitivity improves rapidly from 4 months of age, decreasing by a factor of two until adolescence (Knoblauch et al., 2001). Other research has investigated infant colour perception, for example, there is evidence that infants have at least rudimentary levels of colour constancy at 3-4 months (Dannemiller, 1989; Yang et al., 2015), that is still developing at 3-4 years old and associated with language (Rogers et al., 2020; Witzel et al., 2021). Investigating infant colour perception has also proved useful for understanding the origins of perceptual phenomena such as colour categorisation and colour preference (Bornstein et al., 1976; Catherwood et al., 1990a; Franklin & Davies, 2004; Skelton et al., 2017). For example, a number of studies have measured how long infants look at different hues, finding a peak in looking time at bluish colours and least looking at yellow-green (Adams, 1987; Bornstein, 1975; Franklin et al., 2008, 2010; Skelton & Franklin, 2020; Teller et al., 2004; Zemach & Teller, 2007). These infant visual preferences for single colours also closely resemble adult preferences when asked to rate how much they like colours (Hurlbert & Ling, 2007; Palmer & Schloss, 2010; Skelton & Franklin, 2020). For example, Skelton & Franklin (2020) found that there was a

significant robust correlation between adult colour preference and how long infants looked at colours. Infants also experience colour illusions (Yang et al., 2010), look longer at appropriately coloured objects than inappropriately coloured ones at 6-8 months (Kimura et al., 2010), and have recognition memory for colour from at least 4 months old (Bornstein, 1976; Catherwood et al., 1987, 1990b).

As outlined previously, both adult sensitivity to colour and adult aesthetic ratings of abstract Mondrian images relate to the distribution of chromaticities and the illumination of natural scenes (Bosten et al., 2015; Burton & Moorhead, 1987; Juricevic et al., 2010; Krauskopf & Gegenfurtner, 1992). However, the timescale of the alignment of adult colour perception with the chromatic statistics of natural scenes is unclear. The chromatic sensitivity to scene statistics may potentially be innate and/or occur early in infancy when the visual system is developing and tuning into the natural visual environment. Mature colour perception may also calibrate to the environment in an adaptive process depending on the type of environment adults are exposed to, for example through colour adaptation (Webster, 2011; Webster et al., 2007; Webster & Mollon, 1997). One way of identifying the contribution of different timescales is to take a developmental approach and to test infants. To date, only one study has explored this and found that even at 4-6 months infants' discrimination ellipse is also oriented along the natural blue-yellow daylight axis, with poorest sensitivity to colours that correspond to the blues and orange-yellows commonly

found in natural scenes (Skelton et al., 2021). One interpretation of this finding is that the early alignment of colour discrimination with chromatic scene statistics is further support for an innate evolutionary timescale. However, given the amount of plasticity in infant perception and the extent to which visual development draws on experience (Blakemore & Cooper, 1970; Emberson, 2017; Nelson et al., 2012), it is also possible that infant colour perception tunes into the natural visual environment during the rapid development of colour vision in the first 3-4 months of life. In support of this theory of developmental tuning, Laeng et al. (2007) found differences in adult colour discrimination that were dependent on the adult's latitude of birth and season of birth. For individuals who live above the Artic Circle in Norway, there are extreme seasonal daylight variations during the winter where there is very little natural daylight. It was found that adults born above the Artic Circle had better colour discrimination for purple-ish hues, and adults born below the Arctic Circle had poorer colour discrimination for greenish hues. Interestingly, adults born in the autumn and winter showed the greatest number of errors at discriminating colours and adults born in the summer were better at discriminating purple-ish hues compared to other adults born in other seasons. Hence, this finding suggests that the natural visual environment may play a role in colour perception. Although developmental tuning of colour perception to the visual environment is an exciting theory, to date there are only two existing studies supporting this idea (Laeng et al., 2007; Skelton et al., 2021), and therefore it is still only a tentative hypothesis. If infant colour perception is optimally tuned to the environment, the alignment of colour perception to the natural blue-yellow

daylight axis may be found for other measures of infant colour perception. As outlined above, adults' aesthetic response also appears to be related to the blue-yellow daylight axis, as adults' preference for chromatic Mondrian images peaks when the chromatic distribution aligns with this axis (Juricevic et al., 2010). Although there is no direct measure of infants' aesthetic response (it is unclear if infants even have one), infants do have visual preferences, look longer at some stimuli than others (Fantz, 1961; Fantz & Miranda, 1975; Fantz & Nevis, 1967), and also look longer at stimuli that adults also find more aesthetically pleasing (Damon et al., 2017; Quinn et al., 2008). Papers 3 and 4 of this thesis therefore investigate whether early visual preferences can be predicted by the extent to which the chromatic distribution of an abstract image falls along the natural blue-yellow daylight axis. In addition, these papers consider the role of other perceptual colour features, such as hue, luminance and saturation in infants' visual preference for abstract images, in order to contribute to a greater understanding of the dimensions that govern how infants perceive colour.

Perceptual development and other visual features

Beyond colour, this thesis addresses questions about other types of visual features and addresses broader theoretical issues in the field of perceptual development. First, in addition to investigating the role of colour in infants' visual preference for abstract patterns, Paper 3 considers the relative contribution of other visual features. The classic infant experiments in the 1950s to 1970s

demonstrated that infants preferentially look at faces and have distinct pattern preferences, such as preferring curved shapes compared to straight edges, preferring bulls-eye patterns compared to horizontal stripes and preferring patterns with vertical symmetry (Bornstein et al., 1981; Bornstein & Krinsky, 1985; Fantz, 1958, 1961, 1963, 1965; Fantz & Fagan, 1975; Fantz & Miranda, 1975; Fantz & Nevis, 1967). This pioneering work initiated interest in the development of infant visual processing and changed the dominant view at the time that infants' visual world was a "blooming, buzzing confusion" (James, 1890). Research on infant visual perception has since revealed remarkable perceptual abilities in infants. For example, even foetuses respond to face-like stimuli (Reid et al., 2017), newborns can preferentially follow a face (Cassia et al., 2004; Farroni et al., 2005; Johnson et al., 1991; Simion et al., 2001) and at 2 days old newborns look longer at their mother's face than a stranger's (T. Field et al., 1984; Pascalis et al., 1995). Furthermore, research in infant visual perception has also shown that from 3 months old, infants become sensitive to pop-out, which is when attention is captured by a certain area of a visual stimulus compared to the dissimilar surrounding stimulus; for instance, attending to odd-one-out features (Adler et al., 1998; Adler & Orprecio, 2006; Catherwood et al., 1996; Colombo et al., 1995; Gerhardstein et al., 1999; Quinn & Bhatt, 1998; Rieth & Sireteanu, 1994; Rovee-Collier et al., 1992; Salapatek, 1975). Further visual attention mechanisms develop in the first year of life, as from 4 months old, infants develop the ability to selectively allocate, shift and sustain attention to stimuli, as well as develop visual attentional control (Atkinson & Braddick, 2012; Colombo, 2001; Hood & Atkinson, 1993; Kulke et

al., 2017). From 3-4 months of age, infants show the ability to engage in bottom-up processing (using low-level stimulus features to guide attention) and top-down processing (using prior knowledge to guide attention) when attending to stimuli (Emberson, 2017). These attentional mechanisms change during development as children show an increasing reliance on top-down processes with increasing age, as they gain knowledge about the world which influences where they prioritise their attention (Açık et al., 2010; Amso et al., 2014; Franchak et al., 2016; Frank et al., 2014; Quinn & Bhatt, 1998; Tummeltshammer & Amso, 2018; Walker et al., 2017). Paper 3 investigates visual preference for abstract images that incorporate a range of visual features, such as faces, odd-one-out features and abstract shapes, and enables a test of the relative importance of these features for infant and toddler looking when presented as part of an abstract image. The goal here is also to identify whether principles of visual preference identified in previous developmental research (e.g., infants prefer faces) apply when incorporated into the design of abstract patterns. If so, these could provide heuristics for designers who aim to design patterns that are optimised for the perception of infants and young children.

Infants also appear to have remarkable sensitivity to statistical regularities. For example, the concept of statistical learning in infancy is becoming more prominent in developmental science (Saffran, 2020; Saffran & Kirkham, 2018), with existing research demonstrating that infants are sensitive to statistical regularities across different domains including auditory language (Pelucchi et

al., 2009; Saffran et al., 1996; L. Smith & Yu, 2008) and visual domains (Bulf et al., 2011; Fiser & Aslin, 2002; Kirkham et al., 2002; Poli et al., 2020; Tummeltshammer & Kirkham, 2013). As outlined above, natural scenes also have statistical regularities, not just in colour but for other visual features of scenes such as spatial features (Burton & Moorhead, 1987; D. Field & Brady, 1997; Hansen & Hess, 2006; Tolhurst et al., 1992; Torralba & Oliva, 2003). There has been some consideration of when the sensitivity to these natural spatial scene statistics develops, with studies claiming that infants show a neural sensitivity to natural texture scene statistics from 9 months old which continues to develop in middle childhood (Balas et al., 2018, 2020), and the natural amplitude spectral slope (another spatial image statistic further explained in Paper 3 and 4) does not appear to be sensitive until approximately 10 years old (Ellemberg et al., 2012). Paper 3 further investigates the role of a range of image statistics in infant and toddler visual preference for abstract patterns in order to further understand what image statistics the visual system is able to compute early in life. Investigating this will potentially contribute to our understanding of perceptual development, provide further clarity on whether the efficient coding hypothesis applies to perception early on in development. The efficient coding hypothesis suggests that the processing of sensory information in the brain is reduced and maximised for efficiency (Barlow, 1961). For example, the visual system cannot process all its visual input, and thus when visual information is processed from the retina to the visual cortex, this visual information is compressed to maximise capacity for the neural coding of visual information in the visual cortex. As a result, it is argued the visual system is

optimised to efficiently represent images of the visual environment. Hence, when viewing natural scenes, the adult visual system is theorised to prefer natural scenes as they contain statistical regularities the visual system is optimised to efficiently encode (Olshausen & Field, 2000; Redies, 2007; Simoncelli, 2003; Simoncelli & Olshausen, 2001). Adult research has also shown a relationship between aesthetics and natural statistical regularities (Juricevic et al., 2010), but this has not been explored in infants and young children. Hence, Papers 3 and 4 also contribute to further understanding of the origins of developmental aesthetics.

Interim summary

The above section of this overview chapter has introduced the relevant literature, concepts and theories to the thesis. This introduction also identified a number of questions, broadly on the topic of colour perception in infants and children, that the four papers of this thesis address. Paper 1 and 2 explores CVD in children. Paper 1 develops and evaluates a diagnostic gamified tablet-based CVD test that can be used to diagnose CVD in young children. Paper 2 takes this further by investigating the impact of CVD on older children and adolescents' education and wellbeing. Paper 3 and 4 explores colour and visual perception in infancy and early childhood and its role in visual preference and developmental aesthetics. Paper 3 investigates whether pattern characteristics and low-level features of professionally designed patterns influence visual preference in infants and toddlers. Paper 4 explores whether infants show a

visual preference for abstract images with natural chromatic scene statistics. The following section of the overview chapter will provide overviews of the rationale, design and findings of these papers.

1.2 Thesis overview and research questions

Paper 1: ColourSpot, a novel gamified tablet-based test for accurate diagnosis of colour vision deficiency in young children

Paper 1 (Tang et al., 2021) presents and validates *ColourSpot*, a new, self-administered, gamified and colour-calibrated tablet-based app, which diagnoses CVD from 4 years old. As outlined in the introduction section above, there is a need for a straightforward, accessible and accurate paediatric test for CVD. Although there are several tests which have been used to detect CVD in children (e.g., the Ishihara test for Unlettered Persons, (Ishihara & Ishihara, 2016)), they each have several limitations including inaccessibility, unsuitability and unknown validity in children. For example, many tests are inaccessible to parents or educators and require specialised equipment or a trained administrator. Some tests also require an understanding of numbers, orientations, shapes and/or animals which makes them unsuitable for young children and children with additional and special educational needs.

ColourSpot addresses these limitations by providing an easily accessible, self-administered and gamified test that is tailored to children as young as 4 years

old. *ColourSpot* is a short interactive game that features animated animal characters and takes less than five minutes to complete. On each trial, there are three targets (that lie on protan, deutan and tritan confusion lines) and eight luminance distractors. Each target type has its own individual staircase which starts at the highest available saturation level. The animated characters tell children to “tap a coloured spot”. Depending on whether a target or a distractor is tapped, the saturations of the targets are altered along the dichromatic confusion lines according to an adaptive staircase procedure. See Figure 1.2 for an example of how a trial may be perceived for an individual with normal colour vision and an individual with CVD.

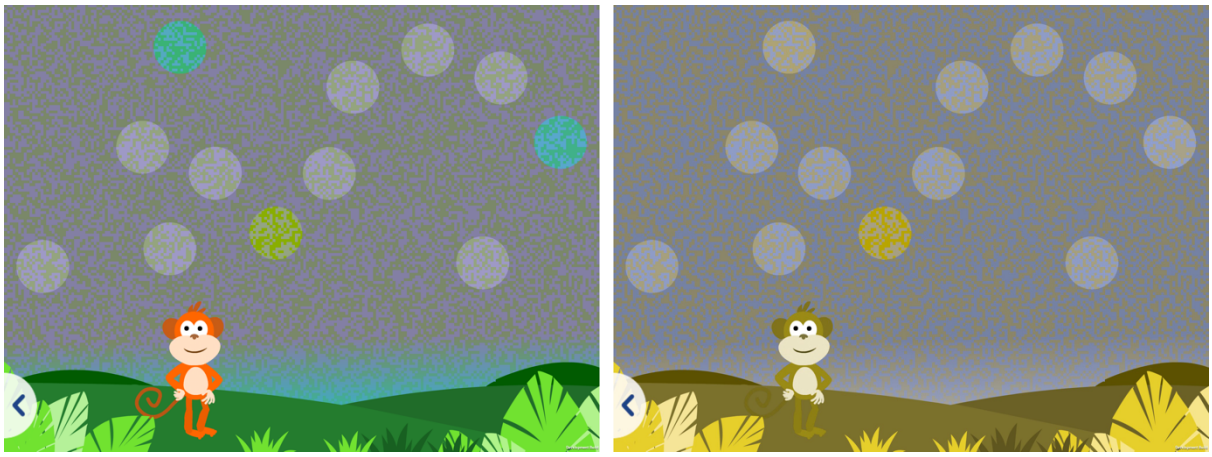


Figure 1.2. A simulation of how an individual with normal colour vision may perceive ColourSpot (left) compared to how a deuteranope may perceive the game (right). Please note that the simulation is for demonstration purposes only and perception may vary amongst CVD observers. The simulation was produced using Vischeck (Brettel et al., 1997; Lillo, Álvaro, et al., 2014).

Seven hundred and seventy-two young children aged 4-7 years old were colour vision screened using the Ishihara test for Unlettered Persons (Ishihara Unlettered; Ishihara & Ishihara, 2016) and the Neitz Test of Colour Vision (Neitz Test; (Neitz & Neitz, 2001), two tests both aimed at children. Depending on the child's diagnosis on the Ishihara Unlettered, each child was categorised into three groups: CVD, inconclusive and colour vision normal (CVN). Any child who made one or more errors on the Ishihara Unlettered were tested on *ColourSpot* as well as a random control group, who made no errors on the Ishihara Unlettered. *ColourSpot* was evaluated against diagnoses made from the Ishihara Unlettered test with a discovery cohort and validated in an independent validation cohort. The discovery cohort was used to identify the optimal classification criteria for diagnosing CVD, and a validation cohort was used to assess the validity of *ColourSpot*'s CVD classification criteria obtained from the discovery cohort.

Results were fitted to non-parametric psychometric functions, where an optimal algorithm detected the best estimates to classify CVD and CVN in the discovery cohort. This led to a "minimum threshold ratio", which calculates the lowest value between the ratio of tritan:protan or tritan:deutan thresholds. This uses tritan as a control measure which is an advantage compared to other paediatric CVD tests as it enables non-visual factors (i.e., lack of attention) that may influence performance to be factored in. Findings showed that children in the CVD group had lower *ColourSpot* minimum threshold ratios than the control

group. From the discovery cohort, a minimum threshold ratio of 0.59 was determined to be the diagnostic criteria for CVD, where a minimum threshold ratio of less than 0.59 is indicative of CVD. The same threshold estimates were applied in the independent validation cohort which revealed consistent findings. The validation cohort showed that compared to the Ishihara Unlettered, *ColourSpot* had a sensitivity of 1.00 and a specificity of 0.97 for classifying CVD. ColourSpot was more decisive about how to classify the participants that the Ishihara Unlettered classified as inconclusive.

Paper 1 indicates that *ColourSpot* can diagnose CVD similarly or even better than the Ishihara Unlettered, as it was also able to categorise children with inconclusive diagnoses. Compared to the Ishihara Unlettered, the Neitz Test generated an unacceptably high level of false positives, which questions its validity as a paediatric CVD test. *ColourSpot* is an accurate paediatric CVD test, and it also enables remote CVD testing which is particularly useful during the current COVID-19 pandemic. *ColourSpot* will be made available to anyone with access to an iPad and it can be self-administered in any setting (e.g., the classroom or the home) by anyone without requiring the child to understand numbers or shapes. The technical features of an automatic colour calibration, and psychophysical threshold estimates along three confusion axes also make it more precise compared to other digitised CVD tests. The development and evaluation of *ColourSpot* in Paper 1 has potential to have a large impact by enabling better, cheaper and more efficient screening of CVD in young children

in schools and in the home. The source code, iPad calibrations and animations have also been made open source so that they can support the development of other paediatric gamified optometric tests, and other remote colour studies using the iPad.

Paper 2: The impact of colour vision deficiency in children and adolescents on self-reported wellbeing, educational engagement and ability to complete school tasks

As outlined in the introduction section, there is a lack of clarity on the effect of CVD on children's education, wellbeing and quality of life. Paper 2 assesses the impact of CVD on wellbeing and school engagement in older children and adolescents. As outlined above, adults with CVD commonly report that CVD affects various aspects of their life such as their health, wellbeing and careers (Barry et al., 2017; Chan et al., 2014; B. Cole, 2004, 2007; Cumberland et al., 2004; Steward & Cole, 1989; Stoianov et al., 2019; Tagarelli et al., 2004). Despite these known difficulties in adults with CVD, the impact of CVD on children is unclear (Mehta et al., 2018; Wilkinson, 1992). The few studies that have investigated this topic provide mixed findings, with some studies suggesting that CVD does not have a negative impact on children's wellbeing and education (Cumberland et al., 2004; Lampe et al., 1973; Nithiyaananthan et al., 2020; Suero et al., 2005), and other studies indicating that children with CVD are at risk of social, behavioural and emotional difficulties as well as poorer educational performance (Grassivaro Gallo et al., 1998; Harrington et al.,

2021; Thomas et al., 2018; Thuline, 1964). There are also some reports that children with CVD have more difficulties than CVN children with colour-related school tasks such as identifying colours and charts in mathematics or science, selecting colours in arts, and identifying different teams during sports (Albany-Ward, 2011; Chan et al., 2014; Sullivan, 2011), and the impact of CVD on school activities has only been explored in two studies (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). In addition, no study has ever investigated whether these reported difficulties with educational tasks lead to poorer educational engagement and affect attitude to school more in children with CVD than CVN children.

The study presented in Paper 2 aimed to systematically investigate the impact of CVD in children and adolescents, and to also quantify the extent that colour-related school tasks are challenging for children with CVD. Forty-five 11–16 year old boys were colour vision screened online in their homes using *ColourSpot*, the remote diagnostic CVD test presented in Paper 1. Each child completed three self-reported online questionnaires measuring wellbeing (KIDSCREEN-10 (Ravens-Sieberer et al., 2010)), educational engagement (Engagement vs Disaffection with Learning (Skinner et al., 2008); Agentic Engagement Scale (Reeve, 2013)), and an additional colour-related tasks difficulty questionnaire that was designed for the study, which rated how difficult they found a list of colour-related tasks in school.

Based on *ColourSpot*'s diagnosis, 29% of children were CVD, and 71% of children were CVN. We recruited partly from online CVD forums which explains the higher rate of CVD than in the typical population. Analyses did not reveal significant differences in wellbeing and educational engagement scores between CVD and CVN groups, suggesting that CVD does not appear to negatively impact wellbeing and educational engagement in older children and adolescents. However, Bayesian analyses indicated anecdotal evidence for the null hypothesis, which suggests that we do not have enough data to make firm conclusions that there is no wellbeing or educational engagement differences between the two groups. Children in the CVD group reported significantly greater difficulty in colour-related school tasks compared to children in the CVN group, with firm support for this effect from Bayesian analyses for naming colours and using colour in art, and substantial support for this effect from Bayesian analyses for completing school tasks (science, design technology, maths and English) when they are coloured or require colour. The analyses quantify previous anecdotal reports from individuals with CVD (Albany-Ward, 2011; Chan et al., 2014; Sullivan, 2011; Waggoner, 2017) and further supports findings from two previous studies (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). As *ColourSpot* also measures discrimination thresholds, a general colour discrimination ability was also measured, which was calculated from the average of a protan, deutan and tritan discrimination threshold. Interestingly, findings indicated that general colour discrimination ability was significantly correlated with difficulty with colour-related school tasks, suggesting the better a

child's colour discrimination is (irrespective of CVD), the more confident they are in their ability to complete tasks at school that involve colour.

The findings of the study presented in Paper 2 suggest that children diagnosed with CVD report that they find colour-related school tasks particularly difficult, but based on the data available, this does not appear to strongly impact the amount that they engage with education or their wellbeing. One question for future research is whether being diagnosed as CVD as a child (and therefore potentially understanding any source of difficulty and obtaining appropriate support), provides some protection against the potential negative impact of CVD on wellbeing, education and quality of life. A replication and extension of the study in Paper 2 using a larger sample and including a sample of children who have CVD but are unaware of their condition, would progress our understanding of this issue. Addressing this question is important for determining whether there is a benefit in re-introducing a school CVD screening programme. For now, the findings of the study in Paper 2 suggest that teachers and parents should be aware that children with CVD lack confidence (and perhaps also ability) in school-related tasks that involve colour, and therefore further support for those children, or modifications to teaching materials would be valuable.

Paper 3: The contribution of various pattern characteristics and image statistics to visual preference in infants and toddlers

Paper 3 investigates visual pattern preferences in infants and toddlers and assesses whether certain pattern characteristics and image statistics predict how long infants and toddlers look at a pattern. As outlined in the introduction section, visual preferences in early development have informed us about the development of visual perception (Fantz, 1958, 1961, 1963; Fantz & Fagan, 1975; Fantz & Nevis, 1967). For instance, infants show distinctive preferences for faces (Cassia et al., 2004; Dannemiller & Stephens, 1988; Fantz, 1961; Gliga et al., 2009; Goren et al., 1975), symmetry (Bornstein et al., 1981; Bornstein & Krinsky, 1985; Fisher et al., 1981), curved shapes (Amir et al., 2011; Fantz & Miranda, 1975; Fantz & Nevis, 1967; Jadva et al., 2010), odd-one-out features (Adler & Orprecio, 2006; Catherwood et al., 1996; Gerhardstein et al., 1999; Quinn & Bhatt, 1998; Salapatek, 1975) and have certain colour preferences (Bornstein, 1975; Brown & Lindsey, 2013; Franklin et al., 2008, 2010; Skelton & Franklin, 2020; Teller et al., 2004; Zemach & Teller, 2007). Although there has been a fair amount of research into infant visual preferences, there has been little investigation of the relative contribution of the preferred visual features to how long infants and toddlers look at more complex patterns that contain these features. A few studies have attempted to measure infant preference for art (Cacchione et al., 2011; Göksun et al., 2014; Krentz & Earl, 2013), but the basis for infants' visual preference for art is very unclear.

As outlined previously, another literature has quantified the properties of natural scenes and complex images in terms of their “scene statistics” –the statistical regularities of natural scene images (Geisler, 2008; Olshausen & Field, 2000; Simoncelli, 2003). Although it has been proposed that the visual system is optimised to the low and high-level image statistics of natural scenes (Berman et al., 2014; Burton & Moorhead, 1987; Kardan et al., 2015), there has been little investigation of how natural image statistics contribute to the development of vision. As discussed above, one study suggests that vision does not tune into certain spatial scene statistics until 10 years old (Ellemberg et al., 2012), although another suggests that infants tune into natural texture scene statistics from 9 months (Balas et al., 2018). A recent study has also shown that infants’ sensitivity to hue is aligned with the distribution of colours in natural scenes along the natural blue-yellow daylight axis (Skelton et al., 2021). Whether or not these and other image statistics contribute to visual pattern preferences early in development has never been investigated.

Paper 3 investigates the visual preferences of infants and toddlers for a set of complex and abstract professionally designed patterns. The study seeks confirmation that certain features such as odd-one-out elements and faces are preferred early in development when shown as part of an abstract complex image; and assesses whether these preferences influence how long infants and toddlers look at complex and abstract patterns (e.g., does designing a face in a pattern make infants and toddlers prefer that pattern over others). The study

also quantifies the low and high-level chromatic and spatial image statistics of the patterns and investigates whether these contribute to how long infants and toddlers look at the patterns. Using eye-tracking, the eye-movements of 6–8 month old infants and 2–3 year old toddlers were recorded whilst they were presented with a set of patterns designed by the baby product company, Cosatto Ltd. The set of patterns had a range of pattern characteristics such as variations in colours and abstract shapes, as well as odd-one-out features, schematic faces and animals. Image analysis quantified the low-level and high-level image statistics of the patterns using chromatic and spatial image statistics. Analyses compared looking time across patterns that contained different features, and also investigated whether the image statistics predicted infant and toddler looking times.

The analyses indicated that toddlers looked significantly longer than infants at the patterns overall. However, infants spent a similar time looking at faces and odd-one-out patterns as toddlers, but less than toddlers at patterns with animals, trees and abstract shapes. One interpretation of these developmental differences is that the detection of faces and odd-one-out features in patterns is a more automatic, bottom-up process that contributes to infant visual preference, but that preference for other features such as animals, trees and certain abstract shapes might rely on more attentional, conceptual and top-down processes that are less mature in infants than toddlers, reflecting age-related attentional changes in bottom-up and top-down processing. The

analyses also indicated that an image statistic, the dominant colour of a pattern, also contributed to visual preference in infants and 2–3 year old toddlers. However, analyses failed to reveal a contribution of the other chromatic and spatial image statistics, although Bayesian analyses revealed there is only anecdotal support for the null hypothesis, indicating more research is needed to make firm conclusions.

The findings of Paper 3 suggest that the presence of certain characteristics in patterns (e.g., faces or odd-one-out elements) will boost infants' preference for those patterns relative to patterns that lack those characteristics. The findings also suggest that the dominant colour of a pattern contributes to how long infants and toddlers look at it. These findings have implications for the design of patterns that are optimised for the visual perception of infants and young children. They also further contribute to our understanding of visual development. Although the dominant colour was the only image statistic to predict infant looking, we are not confident that other image statistics do not also have an effect. Further research that applies our image statistics approach and analyses but uses natural images, where there will be a greater range of scene statistics, may reveal further effects. This further research would be valuable in clarifying the role of natural image statistics in visual development.

Paper 4: Infants do not look longer at abstract images with natural chromatic scene statistics

As mentioned above, it has been argued that the adult visual system is optimally tuned to the image statistics of natural scenes (Simoncelli, 2003; Simoncelli & Olshausen, 2001), and natural scenes show a distinctive chromatic distribution that varies predominantly along the natural blue-yellow daylight axis in colour space (Burton & Moorhead, 1987; Juricevic et al., 2010; McDermott & Webster, 2012; Ruderman et al., 1998). Adults and infants appear to be less sensitive at discriminating colours along this natural blue-yellow daylight axis (Bosten et al., 2015; Krauskopf & Gegenfurtner, 1992; Pearce et al., 2014; Skelton et al., 2021). Adults also rate abstract Mondrian images with chromatic distributions that align with the natural blue-yellow daylight axis as more aesthetic (Juricevic et al., 2010), suggesting that the chromatic naturalness of abstract patterns affects whether adults like them. Although there is evidence that infant hue sensitivity is aligned with the natural blue-yellow daylight axis, it is unclear whether this image statistic also contributes to infants' visual preferences for complex images in a similar way to how it contributes to adults' aesthetic preferences. Evidently, a visual preference is not equivalent to an aesthetic preference in adults (C. Taylor et al., 2013), yet there has been some argument that early visual preferences provide a basis for mature aesthetic preferences to develop (Damon et al., 2017; Massaro et al., 2012; Quinn et al., 2008; Savazzi et al., 2014; Skelton & Franklin, 2020). Therefore, if adults prefer images with natural chromatic distributions, then infants may look longer at them too.

Paper 4 investigates whether infants look longer at abstract images that are more chromatically natural, specifically images with colour distributions that vary closely along the natural blue-yellow daylight axis. Using eye-tracking, 6–8 month old infants were shown chromatic Mondrian images varying along eight colour axes. Following Juricevic et al. (2010), chromatic Mondrian images were created from an algorithm which generated overlapping rectangles of random size, location and varying colour distributions along a single colour axis. Hence, infant looking times were compared for chromatic Mondrian images which varied along the natural blue-yellow daylight axis and other images with an "unnatural" colour distribution. See Figure 1.3 for an example of how the chromatic distribution of the abstract Mondrian images are created.

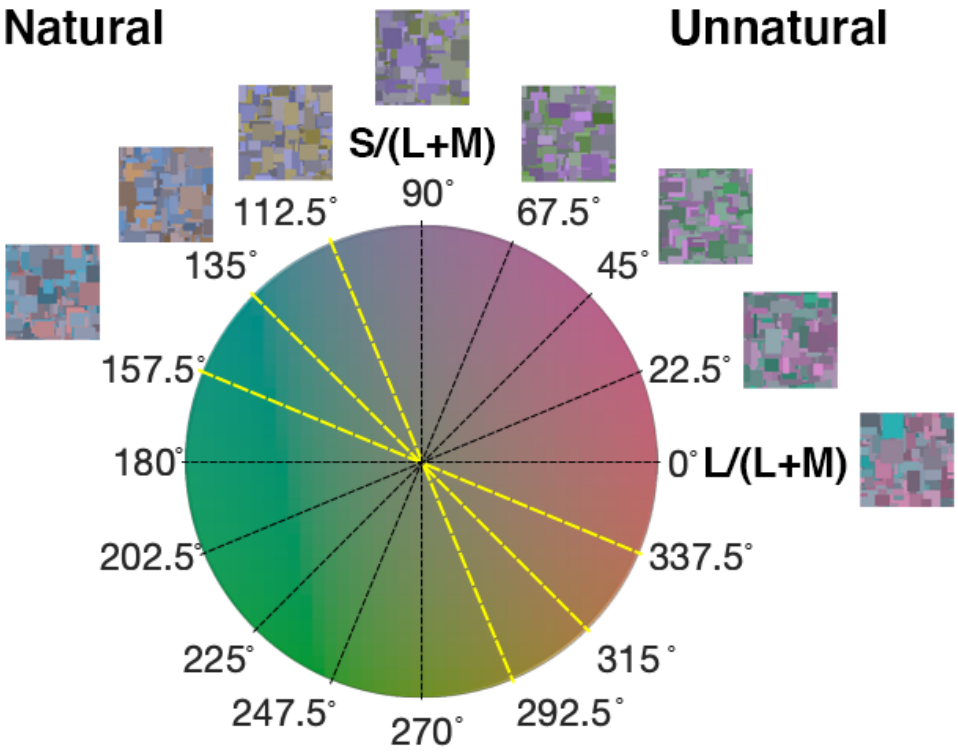


Figure 1.3. Examples of eight types of abstract Mondrian images represented in angles (°) in the MacLeod-Boynton chromaticity diagram. Each image type is created by producing random varying colours of rectangles of random size and locations along a single colour axis in MacLeod-Boynton colour space. For example, the image at 0° consists of colours randomly varying along the 0-180° colour axis. Images along the marked dashed yellow line in colour space (112.5- 292.5°, 135-315° and 157.5- 337.5°) are varying colour distributions closest to the natural blue-yellow daylight axis, and thus would be considered as images with natural colour variation. In contrast, all other images on the opposite colour axes (i.e., at 0-180°, 22.5-202.5°, 45-225°, 67.5-247.5°, 90-270°) would be considered as images with "unnatural" colour variation. Please note that as the size of rectangles were randomised for every trial, a different Mondrian image was produced each time, but each image would have the same chromatic distribution along each colour axis.

The results revealed that there was large variation across individual infants in their preferences for different chromatic distributions, with some infants preferring natural chromatic distributions and other infants preferring Mondrian images with “unnatural” chromatic distributions. However, on average, 6–8 month old infants did not prefer any of the chromatic Mondrian images and looked equally at the images regardless of the naturalness of the chromatic distribution. There was firm evidence for this lack of effect.

The findings of Paper 4 suggest that even though infant hue sensitivity is aligned with the natural chromatic distribution of scenes at 4-6 months old (Skelton et al., 2021), infants do not preferentially look at complex images that have natural chromatic distributions. We discuss a number of interpretations of this lack of effect. First, we consider whether the original finding about infant hue sensitivity aligning to natural chromatic statistics is a false result. Second, we consider hue sensitivity is aligned to the chromatic statistics of the natural environment, but our finding does not translate to infant preferences for complex images. Third, we consider whether the study was sensitive enough and the stimuli appropriate enough to reveal an effect. We suggest further studies that will potentially distinguish between these interpretations. Overall, the study presented in Paper 4 is valuable as it raises the important question of the extent to which infant vision and visual preference tunes into natural chromatic scene statistics and has implications for our understanding of the relationship of infant visual preferences and adult aesthetics.

Chapter 2

Paper 1: ColourSpot, a novel gamified tablet-based test for accurate diagnosis of color vision deficiency in young children

Teresa Tang[†,1], Leticia Álvaro [†,2], James Alvarez[3], John Maule[1], Alice Skelton[1], Anna Franklin[*,1,*] & Jenny Bosten[*,4,*]

[†]Joint first authors

[*]Joint senior authors

[1] The Sussex Colour Group, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

[2] Facultad de Psicología, Universidad Complutense de Madrid, Pozuelo de Alarcón, Madrid 28883, Spain

[3] School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

[4] Sussex Vision Lab, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

2.1 Abstract

There is a need for a straightforward, accessible, and accurate pediatric test for color vision deficiency (CVD). We present and evaluate *ColourSpot,* a self-administered, gamified and color-calibrated tablet-based app, which diagnoses CVD from age 4. Children tap colored targets with saturations that are altered adaptively along the three dichromatic confusion lines. Two cohorts (Total, N=772; Discovery, N=236; Validation, N=536) of 4-7 year old boys were screened using the Ishihara test for Unlettered Persons and the Neitz Test of

Color Vision. *ColourSpot* was evaluated by testing any child who made an error on the Ishihara Unlettered test alongside a randomly selected control group who made no errors. Psychometric functions were fit to the data and “threshold ratios” were calculated as the ratio of tritan:protan or tritan:deutan thresholds. Based on the threshold ratios derived using an optimal fitting procedure that best categorized children in the discovery cohort, *ColourSpot* showed a sensitivity of 1.00 and a specificity of 0.97 for classifying CVD against the Ishihara Unlettered in the independent validation cohort. *ColourSpot* was also able to categorize individuals with ambiguous results on the Ishihara Unlettered. Compared to the Ishihara Unlettered, the Neitz Test generated an unacceptably high level of false positives. *ColourSpot* is an accurate test for CVD, which could be used by anyone to diagnose CVD in children from the start of their education. *ColourSpot* could also have a wider impact: its interface could be adapted for measuring other aspects of children’s visual performance.

Keywords: color vision deficiency, visual development, mobile health applications, pediatric, gamification

2.2 Introduction

Measuring individual visual performance in young children has long been a challenge (Robbins et al., 2003). This has meant that visual problems like color vision deficiency (CVD) or other visual abnormalities are often not detected as

early as they could be. There is a need to develop widely accessible intuitive psychophysical methods targeted to young children.

Congenital CVD is the most common congenital visual disorder, affecting approximately 8% of males and 0.4% of females (Birch, 2012). In many countries there is no routine screening for CVD in schools or by optometrists (Atowa et al., 2019; Azizoğlu et al., 2017; Ciner et al., 1999; Department of Health, 2009; Holroyd & Hall, 1997; Hopkins et al., 2013; Jadhav et al., 2017; Ramachandran et al., 2014; Stewart-Brown & Haslum, 1988; WHO Programme for the Prevention of Blindness and Deafness, 2003). One of the barriers to mass screening for CVD in children is a lack of widely accessible, high quality and child-appropriate diagnostic tests. However, identifying CVD in young children is necessary to support them in accessing an educational system that relies heavily on color.

Normal color vision involves the comparison of signals sent by three photoreceptor types: short (S), medium (M) and long (L) wavelength sensitive cones. Anomalous trichromacy is the more common and milder form of red-green CVD, where all three cone types are present and functional, but there is an abnormality either in the L or in the M cone photopigment's spectral sensitivity. Dichromacy is the more severe form of red-green CVD where there is an absence either of the L or of the M photopigment (Neitz & Neitz, 2000; Parry, 2015; Sharpe et al., 1999). It is important to detect CVD and identify how

individuals can be supported, as it can significantly affect quality of life such as health, wellbeing, and work (Barry et al., 2017; Chan et al., 2014; B. Cole, 2004, 2007; Cumberland et al., 2005; Steward & Cole, 1989; Tagarelli et al., 2004). Ninety percent of adults with CVD encounter problems in their daily lives (Steward & Cole, 1989), while children with CVD are at risk of social, behavioral and emotional difficulties, and adverse educational outcomes (Grassivaro Gallo et al., 1998, 2002; Suero et al., 2005; Thomas et al., 2018; Thuline, 1964). Educational materials, especially those for young children, such as reading schemes or mathematics activities (Birch, 2001), typically rely on color as a learning tool (Rinaldi et al., 2020; Suero et al., 2005), which makes them less accessible for children with CVD. In fact, one simulation of CVD suggests that 10% of tasks in educational textbooks are inaccessible for a child with CVD (Torrents et al., 2011). The dependence on color in education likely means that students with CVD are disadvantaged compared to their peers with normal color vision (Mehta et al., 2018).

Although there are a number of tests which have been used to detect CVD in children, they each have limitations. The gold standard diagnostic test in adults, the anomaloscope, has been used successfully with children older than 7, but the task of mixing a red and green to match a yellow is too demanding for younger children (Verriest, 1982). There are tests which aim to be child-friendly versions of other adult tests for CVD (e.g., the Ishihara test for Unlettered Persons), but their success varies greatly. Table 2.1 summarizes the properties

of existing tests for CVD that have been designed for children or that are presented on tablet displays. An additional table summarizing tests for CVD aimed at adults which have also been used in children is included as Table S1 in the Supplementary Information (SI).

Table 2.1. A summary of tests for CVD either intended for use on children or presented on tablet displays. The table outlines the test type, adult sensitivity, and specificity values with their sample size and comparison test, rates of successful test completion in children with sensitivity and specificity if available, the recommended minimum age for test completion, and the limitations of each test.

Test name	Adult sensitivity/specificity, numbers of adult test participants, and comparison test	Details of tests on children including numbers of participants and sensitivity/specificity where available	Recommended minimum age for test	Limitations
Pseudoisochromatic				
Ishihara for Unlettered Persons* (Ishihara & Ishihara, 1943)	0.98/1.00 (Birch & McKeever, 1993) (CVD=29, CVN=263) Ishihara 1989 edition (Ishihara, 1989)	90% of 3-6 year old children (N=40) successfully completed the test (Mäntyjärvi et al., 2000)	4 years (Birch & Platts, 1993)	†, ‡, § Requires shape knowledge and pathway tracing
Kojima-Matsubara Color Vision Test plates* (Matsubara & Kojima, 1957)	0.08/0.90 (D. Lee et al., 1997) (CVD=13, CVN=20) Anomaloscope	3-6 year old children (N=40) successfully completed the test (Mäntyjärvi et al., 2000)	4 years (Mäntyjärvi et al., 2000)	†, ‡, § Requires animal knowledge
Pease Allen Color Test* (Pease & Allen, 1988)	0.87/1.00 (Pease & Allen, 1988) (CVD=23, CVN=210) Anomaloscope	97% of 3-6 year old children passed the test (Pease & Allen, 1988)	3 years (Pease & Allen, 1988)	†, ‡
Color Vision Testing Made Easy (CVTME)*	0.97/0.90 (Dain, 2010) (CVD=41, CVN=42)	Children over 4 years successfully completed	3 years (Waggoner, 1994)	†, ‡, § Requires shape, animal

(Waggoner, 1994)	Anomaloscope	the test (Richardson et al., 2008)		and object knowledge
Neitz Test of Color Vision* (Neitz & Neitz, 2001)	1.00/0.86 (Block et al., 2004) (CVD=14, CVN=26) Anomaloscope	Tested in 4-12 years (N=115) and verified with genetic testing (Neitz & Neitz, 2001)	4 years (Neitz & Neitz, 2001)	†, ‡, § Requires shape knowledge
Color Vision Evaluation Test (CVET)* (Fish et al., 2020)	N/A	3-18 years 0.96/0.96 (Fish et al., 2020) (CVD=70, CVN=85) Ishihara 38-plate edition	3 years (Fish et al., 2020)	†, § Ability to identify orientations
Oddity				
Mollon-Reffin Minimalist Test* (Mollon et al., 1991)	N/A	3-10 years successfully completed the test (Shute & Westall, 2000). Children rated as most enjoyable test compared to CVTME, Neitz and Analphabetic Ishihara (Tekavčič Pompe & Stirn Kranjc, 2012)	3 years (Shute & Westall, 2000)	†, ‡
University of Waterloo Colored Dot Test* (Hovis et al., 2002)	0.57/1.00 (Hovis et al., 2002) (CVD=21, CVN=31) Anomaloscope	2.5-5 years 0.48/0.97 (Hovis et al., 2002) (CVD=25, CVN=524) Standard Pseudoisochromatic Plates (Ichikawa et al., 1979; Tanabe et al., 1978)	3 years (Hovis et al., 2002)	†
Tablet-based				
DoDo game* (Nguyen, Do, et al., 2014; Nguyen, Lu, et al., 2014)	N/A	6-17 years 0.81/1.00 (Nguyen, Do, et al., 2014) (CVD=16, CVN=16) Ishihara 1998 edition (Ishihara, 1998)	2.5 years (Nguyen, Lu, et al., 2014)	†
Optopad (de Fez, Luque, Matea, et al., 2018)	0.75/0.94 (de Fez, Luque, Matea, et al., 2018) (CVD=16, CVN=50)	3-11 years 1.00/1.00 (de Fez, Luque, Matea, et al., 2018)	3 years (de Fez, Luque, Matea, et al., 2018)	†, § Ability to identify

	Farnsworth Munsell 100-Hue (Farnsworth, 1943)	(CVD=6, CVN=335) Ishihara 2002 edition (Ishihara, 2002)	orientations of the Landolt C

Note. **Definitions. Anomaloscope**: The anomaloscope is an optical instrument where individuals are asked to match different mixtures of red and green monochromatic lights to a yellow monochromatic light. It is the gold standard for assessing color vision; **CVD:** Participants with color vision deficiency (any CVD type, e.g., anomalous trichromacy, dichromacy); **CVN:** Participants with normal color vision; **N/A:** Not available; **Pseudoisochromatic tests**: These tests have an array of colored dots that form a figure (digits, pathways, letters, animals, or shapes) against an isoluminant background which individuals are asked to identify; **Oddity tests**: An odd-one-out task where individuals are asked to identify a colored target amongst distractors; **Sensitivity**: The rate at which a diagnostic test identifies true positives (i.e. individuals with a condition are correctly identified). For example, against the comparison test (in this example, the standard Ishihara test), the Ishihara Unlettered has a sensitivity of 0.98, indicating that 98% of individuals are correctly diagnosed as having a CVD (of any type) and 2% are false negatives (i.e., the Ishihara Unlettered diagnosed the individual as having normal color vision (CVN) but the standard Ishihara test diagnosed the same individual as CVD); **Specificity**: The rate at which a diagnostic test identifies true negatives (i.e. correctly identifies the absence of a condition). For example, when compared with the comparison test (in this example, the standard Ishihara test), the Ishihara Unlettered has a specificity of 1.00, indicating that 100% of CVN individuals were correctly categorized as CVN, and 0% of individuals were false positives (i.e., where the Ishihara Unlettered diagnosed the individual as CVD but the standard Ishihara test diagnosed the same individual as CVN).

Symbols. * The test was specifically designed for children. **† Inaccessibility.** The test is inaccessible for the public and/or requires specialized equipment and/or resources and/or a trained specialist administrator. **‡ Unknown validity**. The sensitivity and specificity values of the tests are unknown in children; **§ Unsuitability.** The test requires an understanding of numbers, orientation, shapes, and/or animals, or the task is so demanding that it is unsuitable for young children and children with additional educational needs.

Although Table 2.1 presents many options for diagnosing CVD in children, no existing test is suitable for mass-screening for CVD from as young as 4 years. Firstly, many of the tests require specialist equipment or resources and/or require a trained administrator. This requires screening of children for CVD to take place in an optometrist's clinic or as part of a well-funded school CVD screening program. The World Health Organization does not currently recommend that color vision screening is included as part of population-based vision assessments (WHO Programme for the Prevention of Blindness and Deafness, 2003), and many countries do not always perform pediatric color vision screenings (Atowa et al., 2019), including Australia (Hopkins et al., 2013), India (Jadhav et al., 2017), Malaysia (Thomas et al., 2018), Turkey (Azizoğlu et al., 2017), the United Kingdom (Department of Health, 2009) and 80% of states in the United States (Ciner et al., 1999). Secondly, although some tests have good sensitivity and specificity in adults, there is little evidence that they can diagnose CVD accurately in young children. Thirdly, many of the tests are not tailored to the capabilities of young children. For example, pseudoisochromatic plate tests may be difficult if children are not yet sufficiently familiar with the shapes, numbers and animals depicted on the plates (Tekavčič Pompe & Stirn Kranjc, 2012). Children may fail to identify these stimuli for reasons that are unrelated to their color vision. Many of the tests rely on the ability to integrate elements into a holistic percept, a skill which is known to be underdeveloped in young children (Kovács, 2000; Scherf et al., 2009). Some tests have

complicated instructions, take too long to complete, or are not sufficiently engaging for children's limited attention and motivation.

One potential solution to the limited accessibility of pediatric tests for CVD is to use tablet-based methods. Such methods can provide mass testing at home, school and in the community, and are colorimetrically adequate for color vision testing if properly calibrated (Bodduluri, Boon, & Dain, 2017; Dain & Almerdef, 2016; de Fez et al., 2016; de Fez, Luque, García-Domene, et al., 2018). Two iPad-based tests for CVD in children, the DoDo game (Nguyen, Do, et al., 2014; Nguyen, Lu, et al., 2014) and the Optopad (de Fez, Luque, Matea, et al., 2018), have been recently developed. Both tests only partially address the limitations of pediatric tests for CVD listed in Table 2.1. The DoDo game has the advantage of gamification (Abramov et al., 1984; Bodduluri, Boon, Ryan, et al., 2017; Ling & Dain, 2018), but its calibration process has not been specified, which means that its effectiveness may vary between devices. Optopad is not gamified or specifically designed for children, and therefore may not sufficiently engage children's motivation and attention, both factors that are known to influence children's task performance (Diez et al., 2001; W. Taylor, 1970). Optopad also requires specialized software and hardware for calibration to be usable on different devices. Furthermore, Optopad identified only 6 out of 341 children as having CVD, a small sample to estimate the test's sensitivity, and a surprisingly small prevalence rate (1.76% compared to an expected prevalence for their sample of 4.69% (Birch, 2012)). This low measured prevalence along

with low adult sensitivity and specificity (0.75/0.94) and the need to exclude some children (younger participants could only enrol after confirming that they were able to read numbers and correctly identify directions) means that Optopad's child sensitivity and specificity values should be interpreted with caution. Both the DoDo game and Optopad are also not yet readily available for download, and there is nothing yet published to suggest that they are being used to diagnose CVD in children.

Here we present *ColourSpot,* a novel gamified tablet-based test for CVD in children from 4 years of age, which overcomes the limitations of existing pediatric tests for CVD. Our aim is for *ColourSpot* to be an accessible mass screening tool for CVD, usable by parents, teachers, or optometrists in the home, classroom, or clinic at the age that children start school. We combine rigorous psychophysics and color calibration with a child-friendly gamified and animated interface. Using an adaptive staircase procedure, *ColourSpot* measures discrimination thresholds for colored targets defined along the three (protan, deutan and tritan) dichromatic color confusion lines (the terms "protan", "deutan" and "tritan" refer to types of CVD relating to the L, M and S cone respectively; protan, deutan and tritan confusion lines are lines in color space along which colors are not discriminable by dichromats missing the respective cone type). Thresholds are measured in an engaging and simple game where children tap the colored targets among luminance-defined distractors to reveal animated characters.

We outline the procedure and associated accuracy for calibrating display devices, and the psychophysics and design of *ColourSpot.* We present results evaluating *ColourSpot* in a two-stage design that ensures that its validation is based on data from a (validation) cohort independent of the (discovery) cohort used to identify optimal classification parameters for CVD versus normal color vision. Two hundred and thirty-six boys in the discovery cohort and 536 boys in the validation cohort completed the Ishihara test for Unlettered Persons (Ishihara Unlettered) and the Neitz Test of Color Vision (Neitz Test). We were interested in evaluating the latter since it has been proposed as a mass-screening tool for use in classrooms. Any child who made more than one error on the Ishihara Unlettered was tested on *ColourSpot*. A randomly selected control group of children who made no errors on the Ishihara Unlettered was also tested on *ColourSpot*. We present *ColourSpot'*s sensitivity and specificity against the Ishihara Unlettered as a pseudo gold standard[1].

[1] The Ishihara Unlettered was used as a pseudo gold standard test as it is the best paediatric colour vision test available as the anomaloscope is not suitable for young children.

2.3 Methods

Development of *ColourSpot*

Color calibration of iPad screens

In order to render colored stimuli accurately on different models of iPad which vary in their display of color, seven models of iPad (iPad 2, 2011; iPad (3rd generation), 2012; iPad (4th Generation), 2012; iPad Air, 2013; iPad Air 2, 2014; iPad Pro 9.7", 2016; and iPad (5th Generation), 2017; Apple, Cupertino, CA, USA) were color calibrated. The gamma functions and the spectral power distributions of the display primaries were measured for a minimum of two units of each model using a PR655 Spectroradiometer (PhotoResearch, Chatsworth, CA, USA). Mean LMS to RGB transformation matrices were created using the radiance spectra of the three primaries, and mean gamma exponents were measured (and corrected). To check the quality of the calibration for each model, we presented our calibrated protan, deutan and tritan stimuli on each iPad model, and measured the spectra of the stimuli, calculating the LMS values of the presented stimuli from the spectra. Calibration was successful, showing relatively small residual systematic errors between the intended and measured chromaticities. The calibration errors tended to be systematic and therefore to place the stimuli on alternative but equally valid confusion lines (see Figure 2.1). *ColourSpot* automatically detects the iPad model in use and provides the relevant iPad calibration for that model in order to achieve colorimetric accuracy for the test.

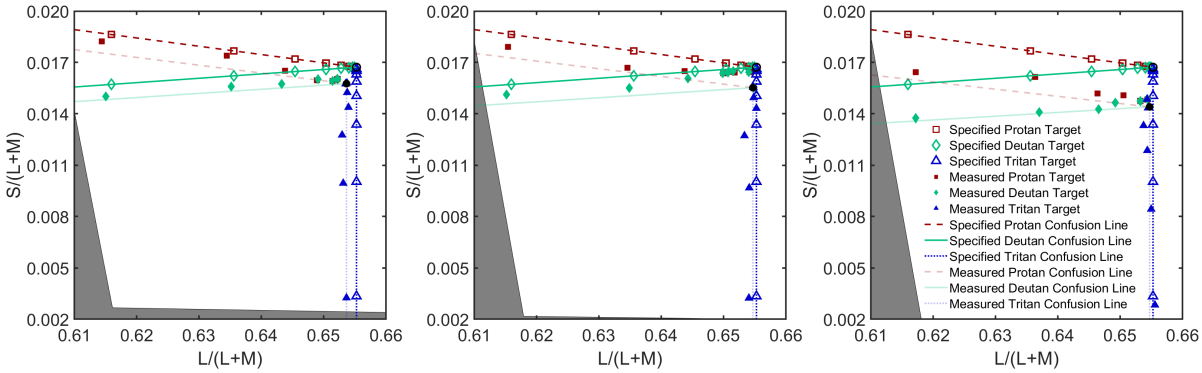


Figure 2.1. Calibration results for iPad Pro 9.7" (left), iPad (5th Generation) (middle), and iPad Air 2 (right) plotted in the MacLeod-Boynton chromaticity diagram (MacLeod & Boynton, 1979). The open symbols show the specified chromaticities and the closed symbols show the measured chromaticities. The red squares are targets on the protan confusion line (dashed line), the green diamonds are targets on the deutan confusion line (solid line) and the blue triangles are the targets on the tritan confusion line (dotted line). Each confusion line passes through the specified (open black circle) or measured (closed black circle) standard D65 white point. Opaque lines represent specified confusion lines (passing through the specified white point) and semi-transparent lines represent the confusion lines closest to the measured stimuli (passing through the measured white point). The solid grey area is outside the iPad's gamut.

Luminance and tritan noise

There are large individual variations in spectral luminosity, particularly for individuals with CVD whose cone spectral sensitivities differ markedly from those of normal trichromats (Judd, 1945; Regan et al., 1994; Teller et al., 2003). To address this, we modeled the maximum luminance signal from our chromatic targets for dichromats and found this to be 10% of the background luminance. We therefore masked this potential signal by drawing the luminance of the targets and distractors from a linear distribution between ± 20% of the average

luminance. Similarly, due to individual differences in cone fundamentals (Bosten, 2019) and small residual color calibration errors, binary tritan noise was added over the stimulus display at ± 16% of the background S/(L+M) value. This value was chosen after estimating the maximum available tritan signal from the protan and deutan targets (due either to individual differences in the peak sensitivities of the cone fundamentals or to calibration error) at 8%.

ColourSpot Design

ColourSpot measures protan, deutan and tritan discrimination thresholds. On each trial, participants are shown three targets selected along each of the dichromatic confusion axes (one protan, one deutan and one tritan target, see Figure 2.1), and eight achromatic distractors varying in luminance (see Figure 2.2). The test begins with a tutorial animation where an animated monkey character jumps up and touches a colored target whilst a child's voiceover says, 'tap a colored spot'. When a colored target is tapped it disappears to reveal a smaller animated character that moves and makes an appropriate animal sound (see Figure 2.3). The tutorial uses highly saturated colored targets that are visible for both individuals with normal color vision and individuals with CVD. After the monkey demonstrates the rule of the game three times, the child begins practice trials. If the child taps a colored spot successfully five consecutive times in the practice trials, they progress automatically to the main part of the test which measures thresholds for the protan, deutan and tritan targets. On each trial, if a target is tapped, a cartoon animal animation along

with an associated sound (e.g., a bird and a chirp) is revealed as a reward. There is no sound or animation if a distractor is tapped. Periodically, child voiceover clips are presented saying “Well done!”, “Amazing” and “You’re good at this!” as positive reinforcement for engaging with the game. During the main part of the test there are 5 animated scene changes to keep the children engaged (Figures 2.2-2.3 and supplementary video S1).

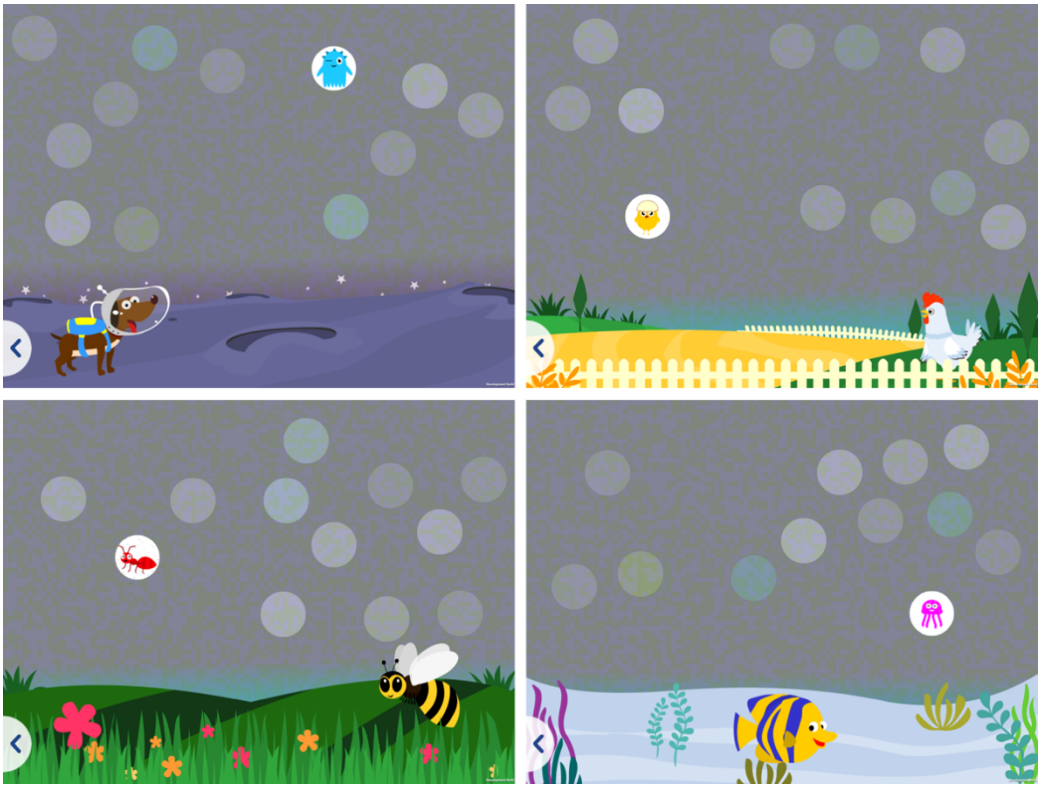


Figure 2.2. Examples of the various scenes showing an animation revealed when a target stimulus is tapped.

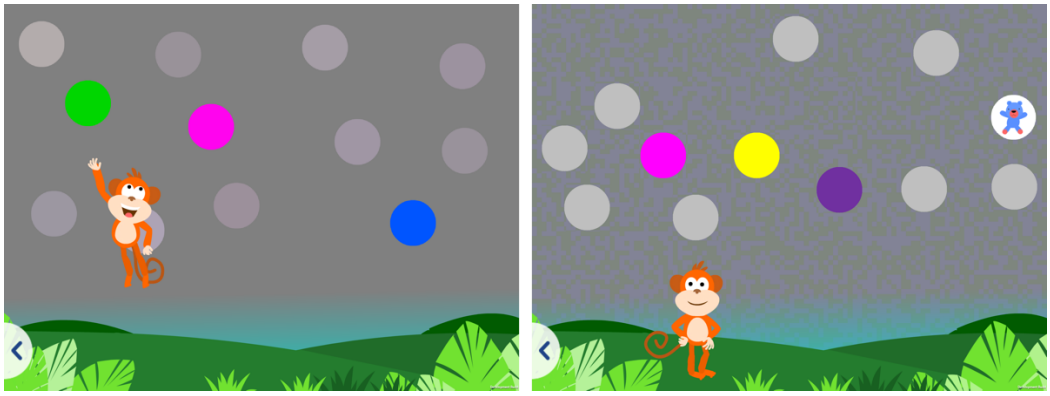


Figure 2.3. (Left**)** A demonstration tutorial where the monkey character demonstrates the rules of the game by tapping the highly saturated practice targets. **(**Right**)** A practice trial example where the cartoon animal animation is revealed from behind the colored practice target when the target is successfully tapped.

To measure thresholds, the colored targets in the main part of the test vary along the protan, deutan and tritan confusion lines according to an adaptive staircase procedure. Each target type (protan, deutan, tritan) has its individual staircase which begins at the highest available saturation level. If a target of a particular type is tapped, its saturation is multiplied by a factor of 0.5 for the next trial, thus decreasing the saturation of targets of that type, making detection of the target type more difficult. If a distractor is tapped, the saturations of all three target types are multiplied by a factor of 1.5 for the next trial, thus making the next trial easier. This design is intended to be helpful for children with severe CVD who may be performing at floor for protan and deutan targets[2]. Having a

[2] Note that even children with CVD will be able to tap a coloured spot successfully on *ColourSpot* and play the game because there are three target types, and their CVD will only make it relatively harder to

tritan target available should reduce their discouragement when they are unable to see the other targets (see Figure 2.4 for an example trial simulated for a CVD observer). *ColourSpot* has two sequential sets of three staircases, with the first and then second set ending after the participant has reached 35 trials on each of the protan, deutan and tritan staircases. Once a staircase for a target type has reached 35 trials, the target type is no longer shown but is replaced by a distractor. Each trial remains on the screen until a target or distractor is tapped, and the locations of the targets and distractors are varied randomly on each trial. Children complete the test at their own pace, and at a moderate speed it can be completed in under five minutes. Thresholds are computed by fitting psychometric functions (see Results).

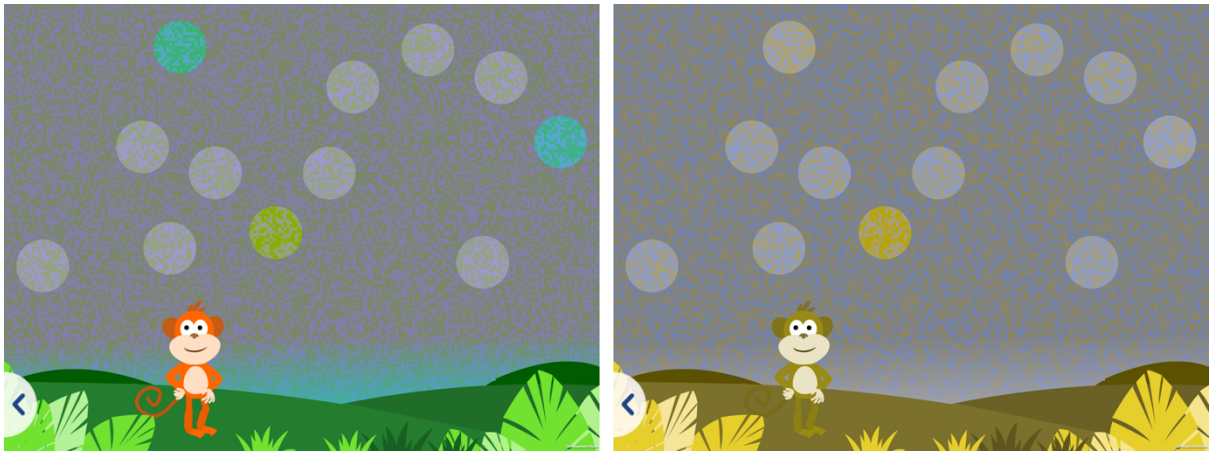


Figure 2.4. A simulation of how a deuteranope may perceive the game (right) compared to a normal trichromat (left). Please note that the simulation is for demonstration purposes only and

see one or two of the target types. This game was designed in this way so that all children would be able to succeed on the task.

perception may vary amongst CVD observers. The simulation was produced using Vischeck (Brettel et al., 1997; Lillo, Álvaro, et al., 2014).

Identifying and Validating *ColourSpot*'s Classification Criteria for CVD

Participants

A total of 772 boys aged 4-7 years (mean 6.11 years, SD 0.90) were recruited from 27 primary schools in Sussex and London, UK. Each school provided information sheets and consent forms to parents inviting their children to participate in the research. Only participants with a completed and signed parent/guardian consent form were permitted to participate in the study. Participants were divided into two independent cohorts. The discovery cohort (N=236, mean 6.17 years, SD 0.83) was used to identify the optimal criteria for assigning CVD versus normal color vision based on the data returned by *ColourSpot*. The validation cohort (N=536, mean 6.09 years, SD 1.04) was used to test *ColourSpot*'s performance as a diagnostic test for CVD, applying the classification criteria identified using the discovery cohort's data. We chose the sample size for the validation cohort using a custom bootstrap method which indicated that with a sample of 120 control children and 40 children with CVD we would have 80% power to identify a 99% sensitivity within an 8% range and a 94% specificity within a 12% range. Based on a prevalence of 8%, a sample of 536 boys would be expected to include 43 boys with CVD. The values for sensitivity and specificity are based on those of the Ishihara test for adults since we were aiming to produce a similarly performing test for children.

Table 2.2 provides a breakdown of the numbers of children of each age in the two cohorts.

Table 2.2. The numbers of participants of each age in the discovery cohort and in the validation cohort.

Age (years)	Discovery Cohort	Validation Cohort
4	17	91
5	95	165
6	75	162
7	49	118
Total	**236**	**536**

The study adhered to the World Medical Association's Declaration of Helsinki (2013), with the exception that it was not pre-registered. Ethical approval was granted by the University of Sussex Science & Technology Cross-Schools Research Ethics Committee (ER/TT283/2) and by the European Research Council Executive Agency.

Color Vision Tests

All children were administered the Ishihara Unlettered and the Neitz Test. The Ishihara Unlettered (Ishihara & Ishihara, 2016) contains a total of eight plates: three are example plates and five are test plates including two plates that involve identifying a geometric shape and three plates that involve curve tracing. The Neitz Test is a multiple-choice pen and paper task that has one

example plate and eight test plates. Each plate contains the outline of a geometric shape presented against a background of grey dots.

Procedure

Each participant was screened for CVD using the Ishihara Unlettered and, in order for us to compare *ColourSpot*'s results with those of another test designed for mass-screening, they also completed the Neitz Test. Any participant who made an error or traced irregularly on three or more plates of the Ishihara Unlettered were assigned to a "CVD" group. Any participant who made one or two errors or traced irregularly on the Ishihara Unlettered was assigned to an "Inconclusive" group (see Figure S1 in the SI) for the distribution of errors on the Ishihara Unlettered). Any child in the discovery cohort who generated inconclusive results on the Ishihara Unlettered were re-tested at a later date if possible (n=6), and their group was reassigned if their results were conclusive at retest (n=4 were reassigned to the control group). Children assigned to the inconclusive group in the validation cohort were not retested. A randomly selected subset of children who made no errors on the Ishihara Unlettered were assigned to a "Control" group.

The children identified as CVD or inconclusive on the Ishihara Unlettered and the control sample were all tested on *ColourSpot*. Three iPad models (iPad Air 2, 2014; iPad Pro 9.7", 2016; and iPad (5^{th} Generation), 2017) were used during testing (also see "Calibration" in the SI). To achieve sufficient trials to measure

protan, deutan and tritan discrimination thresholds, participants had to complete a minimum of 40 trials. In the discovery cohort, one participant in the inconclusive group was unable to undertake *ColourSpot* due to school time restraints. This participant was unable to be re-tested at a later date as they had left the school. In the validation cohort, one participant in the CVD group who had additional educational needs did not want to play *ColourSpot*. Three participants (one inconclusive, two controls) were excluded from the validation cohort as they had completed an insufficient number of trials. Owing to our selection criteria, the administrator of *ColourSpot* was aware of each child's color vision diagnosis by the Ishihara Unlettered when the child was playing *ColourSpot*. However, *ColourSpot* is self-administered by the child with no interference by the administrator once it has started. Therefore, it would be impossible for the administrator to inadvertently bias the child's performance selectively on one color confusion axis in a way that would influence the CVD diagnosis.

Table 2.3. Numbers of participants who played *ColourSpot* in the discovery and validation cohorts. Participants categorized as "CVD" made three or more errors on the Ishihara Unlettered. Participants categorized as "Inconclusive" made one or two errors on the Ishihara Unlettered. Participants categorized as "Control" were randomly selected from children who made no errors on the Ishihara Unlettered. Participants categorized as "Untested" are children who were categorized as having normal color vision by the Ishihara Unlettered but were not tested on *ColourSpot*.

Ishihara Unlettered Groups	**Discovery Cohort**	**Validation Cohort**

CVD	19	37
Inconclusive	7	19
Control	74	118
Exclusions	0	3
Untested	136	362
Total	**236**	**536**

2.4 Results

Data from the two staircases for each of the protan, deutan and tritan targets were combined and fit using non-parametric psychometric functions using local-linear fitting via the software Model-free (Żychaluk & Foster, 2009). Figure 2.5 shows examples of individual psychometric functions from participants in the CVD, inconclusive and control groups. The figure shows that participants in the control group performed similarly for the protan, deutan and tritan targets, whereas participants categorized as CVD by the Ishihara Unlettered had higher thresholds for the protan and deutan targets compared to the tritan targets.

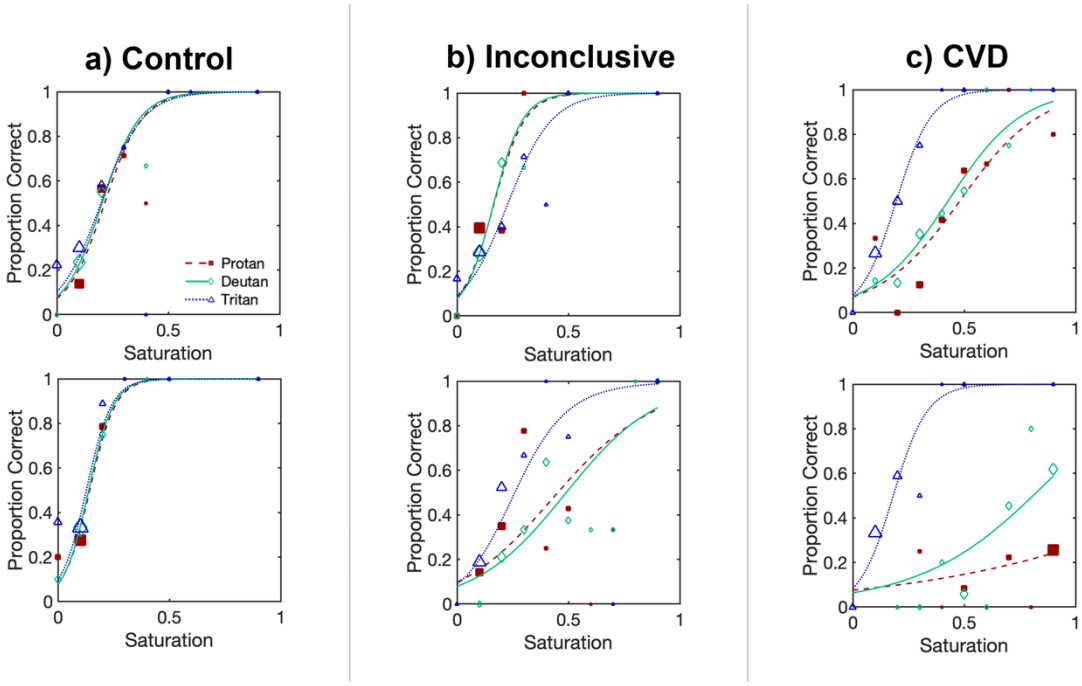


Figure 2.5. Examples of individual psychometric functions for six participants in the Control **(a)**, Inconclusive **(b)** and CVD **(c)** groups fit to their data from *ColourSpot*. Each participant has three psychometric functions representing their performance for detecting protan (red squares, dashed line), deutan (green diamonds, solid line) and tritan (blue triangles, dotted line) targets as a function of *ColourSpot*'s stimulus saturation. The size of the data points is proportional to an arbitrary power of the number of trials of each saturation. *ColourSpot*'s units are relative to the maximum possible saturation of a target on each confusion axis, but the greatest saturation tested on each axis was 0.9 times the maximum. In the chromaticity diagram (MacLeod & Boynton, 1979), the maximum available in gamut saturation (1.0) for protan targets had chromaticity coordinates L/(L+M) = 0.6160, S/(L+M) = 0.0186; the maximum in gamut saturation for deutan targets had chromaticity coordinates L/(L+M) = 0.6160, S/(L+M) = 0.0157; and the maximum in gamut saturation for tritan targets had chromaticity coordinates L/(L+M) = 0.6553, S/(L+M) = 0.0033.

To find the optimal parameters of the Model-free estimation of the psychometric function to achieve the best classification of CVD versus normal color vision

(according to the Ishihara Unlettered) for the discovery cohort, we made a systematic search of parameter space. The optimal parameters were determined as having a fixed bandwidth of 0.70 for the Model-free local linear fit, and a performance level of 0.21 (21%) on which to base threshold estimates (see Figure S2 in the SI). For each participant and pair of fit parameters we found the minimum (most indicative of protan or deutan performance deficit) of the tritan:protan or tritan:deutan threshold ratio. Then we found the pair of fit parameters which maximized the distance between the CVD participant (according to the Ishihara Unlettered) with the largest ratio and the control participant with the smallest ratio. The Model-free estimation was then implemented with the linear algebra library Armadillo (Sanderson & Curtin, 2016, 2018) to be made compatible with Apple software.

Ability of *ColourSpot* to Diagnose CVD

Discovery Cohort

Existing diagnostic tests for CVD are typically based on a criterion score of raw protan and deutan thresholds between individuals (Barbur & Rodriguez-Carmona, 2015; Jurasevska et al., 2014; Mollon et al., 1991; Rabin et al., 2011). We found this to be a suboptimal metric on which to base classification for young children, because the raw thresholds along a given confusion axes do not account for individual differences in non-visual factors such as attention, motivation and engagement that influence task performance (Bosten et al., 2017; Dain & Ling, 2009; Ling & Dain, 2018; see Figure S3 in the SI for a

histogram of our raw protan and deutan thresholds; summary statistics for threshold ratios are provided in Table 2.4). Instead of raw thresholds, we used as a performance metric the minimum value between the ratio of tritan:protan or tritan:deutan thresholds. We propose this metric is effective at factoring out the influence of non-visual factors on task performance as they would affect thresholds along all three confusion axes equally. Given that congenital tritan deficiencies are extremely rare (Wright, 1952), the risk that a high tritan threshold would cancel high protan and deutan thresholds to create a high ratio in the presence of CVD is very low.

Table 2.4. The mean, standard error and 95% confidence intervals of the minimum tritan:protan or tritan:deutan threshold ratios in the discovery and validation cohorts.

Cohort	Ishihara Diagnosis	N	Mean	SE	95% Confidence Intervals	
					Lower Bound	Upper Bound
Discovery	CVD	19	0.23	0.03	0.16	0.29
	Control	74	1.07	0.03	1.00	1.13
	Inconclusive	7	0.83	0.16	0.43	1.23
Validation	CVD	37	0.23	0.03	0.18	0.29
	Control	117	1.08	0.03	1.03	1.14
	Inconclusive	20	0.86	0.08	0.69	1.03

Figure 2.6a shows a histogram of the threshold ratios for the discovery cohort. There is a clear separation using this metric between groups who were classified either as CVD or as having normal color vision (Control) by

performance on the Ishihara Unlettered. Results from the discovery cohort were used to decide on 0.59 as a criterion threshold ratio to define the boundary between a diagnosis of CVD and normal color vision. This was defined as halfway between the maximum threshold ratio for any participant in the CVD group and the minimum threshold ratio for any participant in the control group.

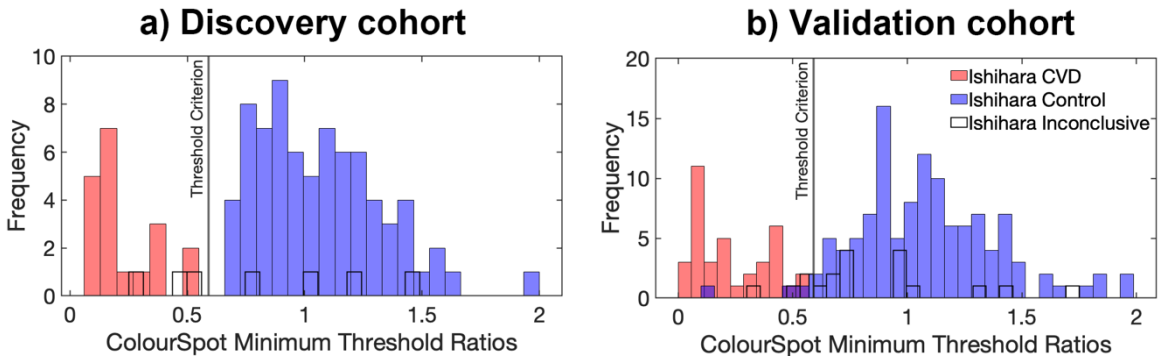


Figure 2.6. Histograms showing, for participants in the discovery cohort **(a)** and the validation cohort **(b)**, *ColourSpot*'s minimum threshold ratios of the tritan:protan or tritan:thresholds for participants categorized as "CVD", "inconclusive" and "control" by the Ishihara Unlettered. The solid vertical line is the criterion threshold ratio used to define the boundary between CVD and normal color vision.

Using the Ishihara Unlettered as a pseudo gold standard test, *ColourSpot* has a nominal sensitivity of 1.00 and a specificity of 1.00 for the discovery cohort (see Table S2 for calculations). Results from *ColourSpot* suggest that the inconclusive group likely includes a combination of participants who have either mild CVD or normal color vision. Figure 2.6a shows that within the inconclusive group, four out of seven participants perform similarly on *ColourSpot* as participants in the control group (and thus likely have normal color vision).

Three inconclusive participants lie within the distribution of threshold ratios for the CVD group (and thus likely have CVD).

Validation Cohort

One hundred and seventy-four participants in the validation cohort played *ColourSpot* (Table 2.3). The optimal fit parameters and the criterion threshold ratio for defining the boundary between CVD and normal color vision from the results for the discovery cohort were applied without modification to the validation cohort (i.e., the classification method and criteria were independently applied to a new sample). Summary statistics for group threshold ratios for the validation cohort are provided in Table 2.4.

Again, using the Ishihara Unlettered as a pseudo gold standard test, *ColourSpot* achieved a sensitivity of 1.00 and a specificity of 0.97 (see Table S3) for the independent validation cohort. If the criterion threshold ratio were applied to the performances of the inconclusive participants, 15 of the 19 inconclusive participants would be classified as having normal color vision, and one as having CVD. Three inconclusive participants were near the boundary between the normal and CVD groups and cannot be classified with confidence by either test. In the control group, three participants performed below the threshold criterion. Two of these were near the criterion threshold ratio and cannot be classified with confidence, but one participant in the control group performed similarly on *ColourSpot* to participants in the CVD group, was likely given a

false negative result by the Ishihara Unlettered (see Figure S4 in the SI and Figure 2.6b).

Neitz Test of Color Vision

A summary of errors on the Neitz Test can be found in Figure S5 and Figure S6 of the SI. Briefly, we found that over 50% of boys made errors consistent with a CVD diagnosis. This implies, in agreement with other studies (Tekavčič Pompe, 2020; Tekavčič Pompe & Stirn Kranjc, 2012), that the Neitz Test does not provide an accurate diagnosis of CVD in young children.

2.5 Discussion

Using gamification and psychophysical methods, *ColourSpot* gathers performance data that can be used to classify CVD similarly to or better than the Ishihara Unlettered. *ColourSpot* has many advantages relative to other pediatric tests for CVD and addresses the limitations of these tests. Firstly, our tritan control stimuli and associated “threshold ratio” measure allow us to distinguish visual from non-visual influences on task performance: If relatively poor performance is specific to the protan or deutan confusion lines, CVD is indicated, but if it applies to all 3 confusion lines, it is likely to be a result of non-visual factors like attention and task engagement. Secondly, gamification with a fun, intuitive and professionally animated interface helps children to maintain engagement and complete the test: 99% of children were able to complete

ColourSpot, including 4 year olds and children with additional educational needs. Thirdly, *ColourSpot* is accessible to anyone with an iPad, allowing it to be administered in any setting. Once downloaded, *ColourSpot* automatically provides the relevant calibration file for each iPad model. It is also self-administered, not requiring a trained administrator. This makes *ColourSpot* highly accessible around the world, either for home use with parents, at school with teachers, in the lab with researchers, or in the clinic with optometrists.

The technical features of *ColourSpot* distinguish it from other digitized tests for CVD. The combination of gamification and an adaptive staircase procedure is a psychophysical method well-suited to children. Particularly for children with CVD, the motivation offered by the game's interface and the fact that stimulus contrast is adapted to current task performance both serve to maintain task engagement by reducing discouragement when they are unable to see targets. The addition of luminance and tritan noise allows us to guard against false negatives that might otherwise be caused by individual differences in color vision, particularly amongst individuals with CVD where cone spectral sensitivities vary significantly from the norm (Judd, 1945; Regan et al., 1994). *ColourSpot*'s tritan control stimuli allow our threshold ratio measure, which confers better classification accuracy for CVD versus normal color vision than using raw protan and deutan thresholds. Automatic application of model-specific color calibration allows mass testing and remote testing in different environments, which is an advantage against other digitized tests for CVD that

are calibrated only for one particular device (Barbur & Rodriguez-Carmona, 2015; Rabin et al., 2011).

Our validation of *ColourSpot* using two independent cohorts provides statistical strength and gives confidence in the accuracy of classification. The classification algorithm optimized for the discovery cohort was applied to the validation cohort without modification, where it demonstrated similarly strong performance for classifying CVD versus normal color vision relative to the Ishihara Unlettered. For the validation cohort the sensitivity (again against the Ishihara Unlettered) was 1.00 and the specificity 0.97. We chose the Ishihara Unlettered pragmatically as a pseudo gold standard because it is not possible to conduct the anomaloscope (the gold standard for adults) on our target age group, owing to its complex task demands. However, the sensitivity and specificity of the Ishihara Unlettered is itself imperfect, evidenced by the fact that it diagnosed 26 children as “inconclusive” in the current study. Using the Ishihara Unlettered as a pseudo gold standard in the current study may therefore lead us to underestimate of the sensitivity and specificity of *ColourSpot* compared to using a fully accurate gold standard if diagnostic errors by the Ishihara Unlettered and *ColourSpot* are independent. Alternatively (and less likely), this may lead us to overestimate sensitivity and specificity for *ColourSpot* if diagnostic errors by *ColourSpot* are correlated with those of the Ishihara Unlettered (for example, if a certain type of mild CVD observer is missed by both tests). It is unclear from the Ishihara Unlettered test alone

whether the 26 participants that scored inconclusively (one or two errors) on the Ishihara Unlettered were anomalous trichromats or made errors on that test for reasons other than CVD (e.g., lapses in task engagement). However, four of the initially inconclusive participants in the discovery cohort were re-tested on the Ishihara Unlettered again within two months and made no errors, suggesting that they constituted false positive diagnoses by the Ishihara Unlettered. In contrast, *ColourSpot* provides a more secure assignment of color vision status for at least some of the inconclusive participants, suggesting that it may be more accurate at diagnosing CVD than the Ishihara Unlettered. In the validation cohort, one participant was diagnosed as normal by the Ishihara Unlettered but was classified as CVD by *ColourSpot*. Given the position of that participant's threshold ratio near the centre of the CVD distribution (see Figure 2.6b and Figure S6 in the SI for the psychometric function of this participant), we believe that this is likely a false negative diagnosis by the Ishihara Unlettered, but a genetic test would be the only definitive way to confirm the diagnosis at the current age of this participant. As the protan and deutan confusion lines are so close to one another in color space, distinguishing protan CVD from deutan CVD is challenging. It is not possible to estimate *ColourSpot*'s accuracy for this classification using our current data as our gold standard test (Ishihara Unlettered) does not accurately distinguish protan CVD and deutan CVD. We plan to explore *ColourSpot*'s ability to classify protan and deutan CVD types by testing it in adults and older children against the anomaloscope.

We aim that accurate diagnosis of CVD using *ColourSpot* from 4 years of age at the start of education will mitigate some of the negative impact of CVD on children's education and wellbeing (Grassivaro Gallo et al., 1998, 2002; Mehta et al., 2018; Suero et al., 2005; Thomas et al., 2018; Thuline, 1964). The *ColourSpot* app provides advice sheets on CVD, designed by Colour Blind Awareness, a non-profit organization to raise awareness of CVD (http://colourblindawareness.org), which aims to help parents and teachers adopt strategies to mitigate some of the potential negative effects of CVD on children. Once generally released, *ColourSpot* can be used by parents, teachers, and optometrists, and we hope that it will enable mass screening for CVD in young children due to its ease of use, psychophysical rigor, and accurate diagnosis.

ColourSpot is available to download for research purposes. The source code can be downloaded at https://osf.io/v5p2y/ and/or a fully compiled version can be made available by contacting the senior authors. As well as the test itself, we also make available for research purposes the interface, animations and general methods which could be applied to measure other visual abilities in children.

Chapter 3

Paper 2: The impact of colour vision deficiency in children and adolescents on self-reported wellbeing, educational engagement and ability to complete school tasks

Teresa Tang & Anna Franklin

3.1 Abstract

Congenital colour vision deficiency (CVD) is a common visual disorder, affecting around one child in every classroom. However, the effect of CVD on children's wellbeing, education and performance in school is unclear. The current study investigates whether CVD is associated with poorer wellbeing and educational engagement in 11–16 year old boys, and aims to quantify the extent to which colour-related school tasks are challenging for children with CVD. Boys were screened for CVD using *ColourSpot*, a remote self-administered diagnostic paediatric CVD test (Tang et al., 2021). Each child also completed three self-report questionnaires measuring wellbeing and educational engagement (KIDSCREEN-10 (Ravens-Sieberer et al., 2010); Engagement vs Disaffection with Learning (Skinner et al., 2008); Agentic Engagement Scale (Reeve, 2013)), and also rated how difficult they found colour-related school tasks. Based on *ColourSpot*'s diagnostic criteria, 13 children were diagnosed as CVD and 32 children were colour vision normal (CVN). Results did not reveal

significant differences in wellbeing and educational engagement scores between CVD and CVN groups, and Bayes factors indicated anecdotal evidence for the null hypothesis. However, the CVD group did report significantly greater difficulty than the CVN group in doing colour-related school tasks. General colour discrimination ability was also significantly related to the amount of self-reported difficulty with colour-related school tasks. The implications for paediatric CVD screening, and for how to support children with CVD in the classroom are discussed.

Keywords: colour vision deficiency, adolescent, wellbeing, educational engagement, telehealth

3.2 Introduction

Congenital colour vision deficiency (CVD) is a common visual disorder where there is an absence or abnormality in one or more of the three cone photoreceptors (L, M, and S cones), affecting approximately 8% of males and 0.4% of females (Birch, 2012). In many countries including the UK, there is no routine screening for CVD in schools or by optometrists (Atowa et al., 2019; Azizoğlu et al., 2017; Ciner et al., 1999; Department of Health, 2009; Holroyd & Hall, 1997; Hopkins et al., 2013; Jadhav et al., 2017; Ramachandran et al., 2014; Stewart-Brown & Haslum, 1988; WHO Programme for the Prevention of Blindness and Deafness, 2003). Many optometrists do not routinely test for

colour vision as it is not mandatory in their assessments (B. Cole, 2007). This has led to many people with an undiagnosed CVD, with 80% of pupils having an undiagnosed CVD when they start secondary school (Albany-Ward, 2005). The importance of screening for CVD in children has been debated (B. Cole, 2015; Long et al., 2015; Ramachandran et al., 2014), with some claiming that because CVD is neither treatable nor progressive it does not meet the screening standard of an important health condition (Logan & Gilmartin, 2004; Ramachandran et al., 2014). Others have argued that there is little evidence to suggest CVD has a negative impact on an individual's quality of life (Cumberland et al., 2004).

However, in fact, there is mounting research to suggest that the impact of CVD on people's lives is often underestimated. Nearly 90% of adults with CVD report encountering problems in their daily life including reading a map, choosing clothes, driving, judging the ripeness of fruit and vegetables, cooking, and selecting colours of materials (Chakrabarti, 2018; B. Cole, 1972; Steward & Cole, 1989; Tagarelli et al., 2004). Those with CVD can also be limited in their career choice as people with CVD are often discriminated or prohibited from working in certain occupations, including armed forces, railways, civil aviation, police service, fire-fighting services, and electrical engineering (Chakrabarti, 2018; Chan et al., 2014; B. Cole, 2007). It is reported that individuals who do not discover they have CVD until their late teens or late twenties often feel grief, anger and disbelief upon discovery of their diagnosis and they realise their

condition may significantly affect their career aspirations which they may have already made an emotional and financial investment pursuing (Long et al., 2015). A CVD-specific questionnaire measuring quality of life in adults has shown that CVD can significantly impact quality of life for health, emotions and careers (Barry et al., 2017). For example, adults with CVD report experiencing negative emotions such as feeling anxious, depressed, unconfident, embarrassed, humiliated and low self-esteem due to issues caused by problems seeing colours. Although Barry et al. (2017) did not find that adults with CVD scored differently on mental health scores compared to adults with normal colour vision, their study highlights the negative emotional impact of having CVD as an adult.

The impact of CVD could potentially be amplified in an educational context as educational materials and teaching tools highly rely on colour coding. For example, colour is used as a teaching and learning tool to engage students to learn new concepts and improve their memory (Dzulkifli & Mustafar, 2013; Rinaldi et al., 2020; Suero et al., 2005). Colour coding is also used as a way for pupils to indicate whether they have understood concepts, and in marking and feedback schemes. Hence, students with CVD are potentially at a disadvantage in the classroom compared to their peers with normal colour vision. One concrete example of this is that 10% of the content in a sample of mathematics textbooks required normal colour vision to answer correctly (Torrents et al., 2011). Students with CVD also often report experiencing difficulties in a variety

of subjects in school, particularly in art, chemistry, biology, physics and mathematics (Albany-Ward, 2011; Chan et al., 2014; Sullivan, 2011; Waggoner, 2017). Some anecdotal examples of the struggles they experience include selecting colours in art; the inability to read litmus paper or differentiate chemical solutions in chemistry; the inability to identify species in plants in biology; problems identifying coloured wiring in physics and difficulties reading pie charts and graphs in mathematics (Albany-Ward, 2011; Chan et al., 2014; Sullivan, 2011; Waggoner, 2017). Confusion and difficulties in sports can also arise in CVD individuals due to the inability to distinguish between the coloured uniforms of both teams and children with CVD may get lost more easily if directions are using coloured objects as signs (Chan et al., 2014; Sullivan, 2011; Waggoner, 2017). The impact of CVD on colour-related tasks has been investigated in primary and secondary school children in South Africa and Nigeria, where children were asked to report frequency of difficulties with colour-related tasks in school and daily life (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). Both studies found that compared to children with normal colour vision, children with CVD report significantly more difficulty with colour-related tasks in school activities such as identifying colours and charts in mathematics, working on the computer, colours in fine arts, identifying colours in crafts and hobbies, and identifying different houses or teams during sports. Additional difficulties with colour-related tasks in daily life for children with CVD include selecting colour of clothes, identifying flowers, judging the ripeness of food, watching sports, describing and recognising cars based on colours and recognising traffic signal lights. Hence, CVD increases difficulties in colour-

related tasks which could potentially make school more challenging for students with CVD compared to students with normal colour vision.

Although it appears that CVD makes certain tasks in school more difficult, whether or not this decreases educational engagement or performance is currently unclear (Mehta et al., 2018). There are inconsistent findings in the literature regarding the impact of CVD on educational outcomes (Wilkinson, 1992). Cumberland et al. (2004) argued that CVD does not have negative educational outcomes, but their study was based on data from a 1958 British birth cohort when colour may not have been as dominant in the classroom as it is today as black and white teaching materials were more standard. Two other studies have failed to find differences in educational performances in young children and primary school-aged children with CVD (Lampe et al., 1973; Suero et al., 2005). However, other studies have reported a negative impact of CVD on children's educational performance. For example, two studies found that children with CVD had lower school achievements compared to children with normal colour vision (Grassivaro Gallo et al., 1998, 2002). It was also recently found that undetected and untreated visual problems in school children, including CVD, was linked with poorer school performance in Ireland (Harrington et al., 2021).

There are also mixed findings on whether CVD is associated with behavioural and emotional issues. A relatively old study argued that students with CVD are

more likely to be referred to the psychologist for behavioural problems than children with normal colour vison during the first years of school (Thuline, 1964). Another study found that primary school-aged pupils with CVD in Malaysia were more likely to have behavioural and emotional problems, such as more internalising problems and social and attentional problems, compared to pupils with normal colour vision, when measured with the Child Behaviour Checklist (CBCL; Thomas et al., 2018). However, there were no differences between CVD and colour vision normal (CVN) Malaysian primary school pupils in behavioural issues when measured with the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997; Nithiyaananthan et al., 2020). This mixed set of findings indicates that the issue of the impact of CVD on children's wellbeing and education warrants further investigation. In particular, it is important to understand the situation for the modern classroom today given the likely increase in usage of coloured materials. Further research also needs to establish whether there is a clear relationship between CVD and any difficulties with wellbeing and educational engagement that can be accounted for by difficulties with colour-related tasks at school.

The current study investigates whether 11–16 year old boys with a diagnosis of CVD, who attend secondary schools predominantly in the UK, self-report having difficulty with colour-related school tasks, and whether their wellbeing and educational engagement scores are lower on reliable and validated self-report questionnaires compared to 11–16 year old boys with normal colour vision. As

the study was conducted during the COVID-19 pandemic, participants took part in the study remotely online on iPads in their own homes. CVD was diagnosed remotely using *ColourSpot* which is a gamified and psychophysical iPad test for CVD that is as accurate as the Ishihara test of Unlettered Persons (Tang et al., 2021). Unlike other paediatric CVD tests, *ColourSpot* measures colour vision on a continuum by measuring chromatic thresholds along the protan, deutan and tritan colour confusion lines. The diagnosis of CVD is dependent on the minimum threshold ratio which calculates the lowest ratio value of the tritan:protan or tritan:deutan threshold, where a minimum threshold ratio of less than 0.59 is indicative of CVD. In addition to playing *ColourSpot*, participants were also asked to complete three validated and reliable self-report questionnaires that measure wellbeing and educational engagement in children and adolescents (KIDSCREEN-10 (Ravens-Sieberer et al., 2010); Engagement vs Disaffection with Learning (Skinner et al., 2008); Agentic Engagement Scale (Reeve, 2013)). Finally, participants were asked to rate how difficult they found specific colour-related tasks in school. These tasks were specified on the basis of anecdotal evidence from discussions with individuals with CVD (Albany-Ward, 2011; Barry et al., 2017; Chan et al., 2014; Waggoner, 2017) and from two studies which have investigated this in children (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). Overall, the study aims to provide further evidence as to whether CVD is associated with wellbeing and educational engagement issues in older children and adolescents, and to quantify the extent to which colour-related tasks in a school context are a challenge for students with CVD.

3.3 Methods

Participants

One-hundred-and-fifty-five people signed up to the study and were recruited online via schools and social media platforms during the COVID-19 pandemic. Participants required access to an iPad to participate. A total of 72 males aged 11-16 years old (*M*= 14.15 years, *SD*= 1.80) participated in this study but 27 participants were excluded from final analyses as they did not complete all tasks, leading to a final sample of 45 males (*M*= 13.82 years, *SD*= 1.85). Only males were recruited due to the higher prevalence of CVD (8%) compared to females (0.4%) (Birch, 2012). Participating countries included: UK (*n*=42), USA (*n*=2) and France (*n*=1). Upon full completion of the study, participants were invited to a prize draw to win one of 30 £20 Amazon vouchers (or an equivalent currency). Ethical approval was granted by the University of Sussex Science & Technology Cross-Schools Research Ethics Committee (ER/TT283/3).

Materials

Colour vision assessment

Colour vision was assessed using *ColourSpot*, a gamified diagnostic tablet-based test for CVD in children (Tang et al., 2021). This test has been validated in 4-7 year old children and has been shown to be an accurate diagnostic test for CVD (Tang et al., 2021). Children tap coloured targets with saturations that are altered adaptively along the three colour confusion lines and the minimum

threshold ratios are calculated from protan, deutan and tritan thresholds. Protan, deutan and tritan relate to types of CVD relating to the L, M and S cones respectively. To obtain colorimetric accuracy, *ColourSpot* automatically detects the iPad model in use and provides the relevant iPad calibration for that model. *ColourSpot* currently has colour calibration files for 9 iPad models including: iPad 2, 2011; iPad (3rd generation), 2012; iPad (4th generation), 2012; iPad Air (1st generation), 2013; iPad Air 2, 2014; iPad Pro 9.7", 2016; iPad (5th generation), 2017; iPad (6th generation), 2018; and iPad (7th generation), 2019. If an unidentifiable or non-calibrated iPad model was found, a default calibration file using the iPad (5th generation) calibration file was provided. Participating iPad models included: iPad Air (1st generation), 2013 (*n*=1); iPad mini 2, 2013 (*n*=1); iPad Air 2, 2014, (*n*=3); iPad Pro 12.9" (1st generation), 2015 (*n*=3); iPad Pro 9.7" (1st generation), 2016 (*n*=2); iPad (5th generation), 2017 (*n*=3); iPad Pro 10.5" (2nd generation), 2017 (*n*=2); iPad Pro 12.9" (3rd generation), 2018 (*n*=1); iPad (6th generation), 2018 (*n*=10); iPad (7th generation), 2019 (*n*=5); iPad (8th generation), 2020 (*n*=2); and unidentifiable iPad models (*n*=12).

Diagnosis of CVD is determined by threshold ratios calculated as the lowest number (minimum) threshold ratio of the tritan:protan or tritan:deutan thresholds. A minimum threshold ratio of less than 0.59 is indicative of CVD. Parental and child self-reports of any previous diagnoses of CVD and additional visual problems by an optician were also recorded. In total, 13 children were reported as already diagnosed as CVD by their parents and 14 children self-

reported as already having a diagnosis of CVD. Additional visual problems such as wearing glasses (*n*=1), myopia (short-sightedness) (*n*=1), hyperopia (long-sightedness) (*n*=2) and exotropia (a type of strabismus) (*n*=1) were reported.

Questionnaires

Wellbeing

The KIDSCREEN-10 is an international 10-item self-reported wellbeing and health-rated quality of life questionnaire for children and adolescents aged from 8-18 years old (Ravens-Sieberer et al., 2005, 2010, 2014; The KIDSCREEN Group Europe., 2006). This questionnaire was selected because it has been shown to be a fast and efficient, valid and reliable measure of wellbeing for the target age group (Erhart et al., 2009; Ravens-Sieberer et al., 2010). Examples of questions asked can be seen in Table 3.1. For each question, participants were asked to rate on a scale of 1 to 5, where 1 was “never/not at all” and 5 was “always/extremely”. A higher total score indicates greater wellbeing.

TABLE 3.1. List of questions from KIDSCREEN-10.

KIDSCREEN-10
1. Have you felt fit and well?
2. Have you felt full of energy?
3. Have you felt sad?
4. Have you felt lonely?

5. Have you had enough time for yourself?
6. Have you been able to do the things that you want to do in your free time?
7. Have your parent(s) treated you fairly?
8. Have you had fun with your friends?
9. Have you got on well at school?
10. Have you been able to pay attention?

Educational engagement

Two questionnaires were used to measure the child's educational engagement including the Engagement vs Disaffection with Learning (Skinner et al., 2008) and the Agentic Engagement Scale (Reeve, 2013). The Engagement vs Disaffection with Learning (EDL) is a 20-item questionnaire measuring a student's emotional and behavioural engagement and disaffection in the classroom, where each item is rated on a 4-point scale, where 1 was "not at all true" and 4 was "very true". The Agentic Engagement Scale (AES) is a brief 5-item questionnaire and was additionally added to measure how students use their initiatives to support their learning, where each item is rated on a 7-point scale where 1 was "strongly disagree" and 7 was "strongly agree". Example questions from both questionnaires can be seen in Table 3.2. A higher total score in both engagement questionnaires would indicate greater educational engagement.

Table 3.2. Questions from two engagement questionnaires including the Engagement vs Disaffection with Learning (EDL) and the Agentic Engagement Scale (AES).

Engagement vs Disaffection with Learning (EDL)	
Behavioural engagement	1. I try hard to do well in school. 2. In class, I work as hard as I can. 3. When I'm in class, I participate in class discussions. 4. I pay attention in class. 5. When I'm in class I listen very carefully.
Emotional engagement	1. When I'm in class, I feel good. 2. When we work on something in class, I feel interested. 3. Class is fun. 4. I enjoy learning new things in class. 5. When we work on something in class, I get involved.
Behavioural disaffection	1. When I'm in class, I just act like I'm working. 2. I don't try very hard at school. 3. In class, I do just enough to get by. 4. When I'm in class, I think about other things. 5. When I'm in class, my mind wanders.
Emotional disaffection	1. When we work on something in class, I feel bored. 2. When I'm in class, I feel worried. 3. When we work on something in class, I feel discouraged. 4. Class is not all that fun for me. 5. When I'm in class, I feel bad.
Agentic Engagement Scale	1. I let my teacher know what I need and want. 2. I let my teacher know what I am interested in. 3. During this class, I express my preferences and opinions. 4. During class, I ask questions to help me learn. 5. When I need something in this class, I'll ask the teacher for it.

Difficulties with colour-related school tasks

A 10-item questionnaire was developed to quantify difficulties with colour-related tasks in school between CVD and CVN children and adolescents. See Table 3.3 for a list of questions. These questions were formed based on anecdotal reports of what individuals with CVD find challenging (Albany-Ward, 2011; Chan et al., 2014; Waggoner, 2017) and also on the basis of the studies which have quantified some of the difficulty with colour-related tasks in school for children with CVD (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). Participants were asked to rate how difficult they found each task on a scale of 5, where 1 was “strongly disagree” and 5 was “strongly agree”. A higher score would indicate greater difficulty with colour-related school tasks.

Table 3.3. List of questions for difficulties with colour-related school tasks.

Difficulties with colour-related school tasks	
I find it difficult to:	
	1. Name the colour of things
	2. Do Maths or English exercises when they are coloured.
	3. Understand coloured diagrams, charts and figures
	4. Do things in science or design technology that involve colour (e.g., reading litmus paper, seeing the colour of a flame, or knowing when meat is cooked)
	5. Use colour in art and be confident in selecting the right colours

	6. Identify who is on my sports team when people are wearing coloured bibs and sports kit.
	7. See the coloured ball when playing sports.
	8. Read letters and words when displayed in colour on the white board or textbooks.
	9. Find my way around the school.
	10. Understand my teacher's feedback when they mark my work with a coloured pen.

Design

The main independent variable was the between-groups categorical variable of colour vision status (CVN vs. CVD). Another independent variable was the continuous variable of *ColourSpot*'s average colour discrimination threshold, which represents how well children discriminate colour generally, irrespective of CVD. The dependent variables are the participant's questionnaire scores on wellbeing (KIDSCREEN-10), educational engagement (EDL and AES), and difficulties with colour-related school tasks.

Procedure

Parents completed the information sheet and consent form online via Qualtrics. After completion of the online information sheet, an automatic email was sent to the parents, providing them with instructions to complete the tasks including a link to complete all the questionnaires on Qualtrics and instructions on how to download *ColourSpot* on their iPad. The study lasted less than 20 minutes was completed independently by participants online on their iPad at home. Parents

were notified if *ColourSpot* indicated that the participant is likely to have CVD and were pointed to an optician for clinical diagnosis and provided with educational advice sheets developed with the charity Colour Blind Awareness (http://colourblindawareness.org).

3.4 Results

Diagnosing CVD with *ColourSpot*

Based on *ColourSpot*'s diagnostic criteria, any participant with a minimum threshold ratio less than 0.59 was categorised as CVD. Children were categorised into CVD (N=13; *M*= 13.77 years, *SD*= 1.82) or CVN (N=32, *M*= 13.83 years, *SD*= 1.89) groups, see Figure 3.1. This led to a sample of 29% of CVD children and 71% CVN children. This is a higher prevalence of CVD compared to the population as we partially recruited from online CVD forums. Figure 3.2 shows examples of individual psychometric functions of four participants in the CVD and CVN group. These psychometric functions have been fitted using a non-parametric psychometric function via Model-free (Żychaluk & Foster, 2009) and show the participant's performance on protan, deutan and tritan targets. We can see from Figure 3.2 those participants in the CVD group have poorer protan and deutan thresholds compared to the CVN group who perform similarly on all three protan, deutan and tritan axes.

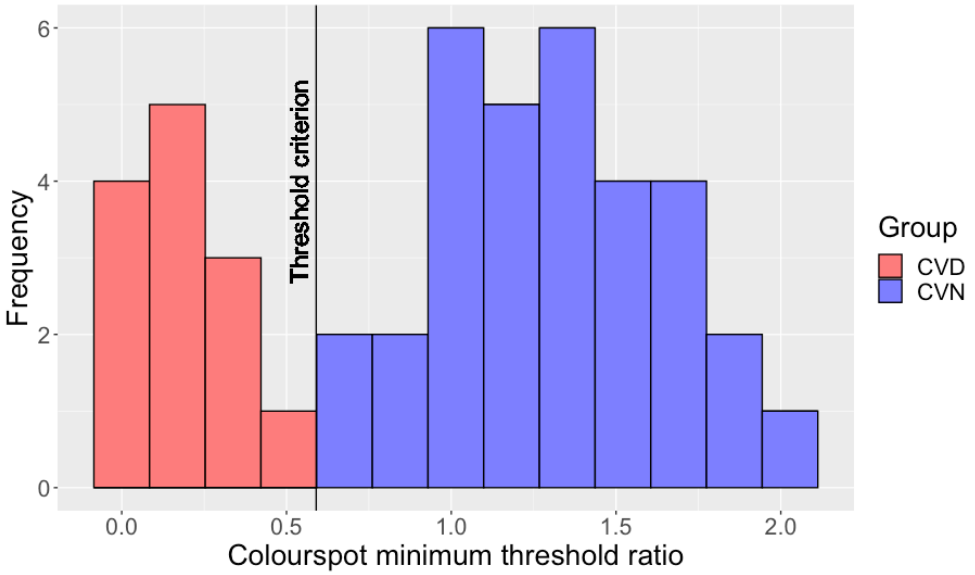


Figure 3.1. Histogram showing the number of participants in the CVD and CVN group as a function of their *ColourSpot* minimum threshold ratio.

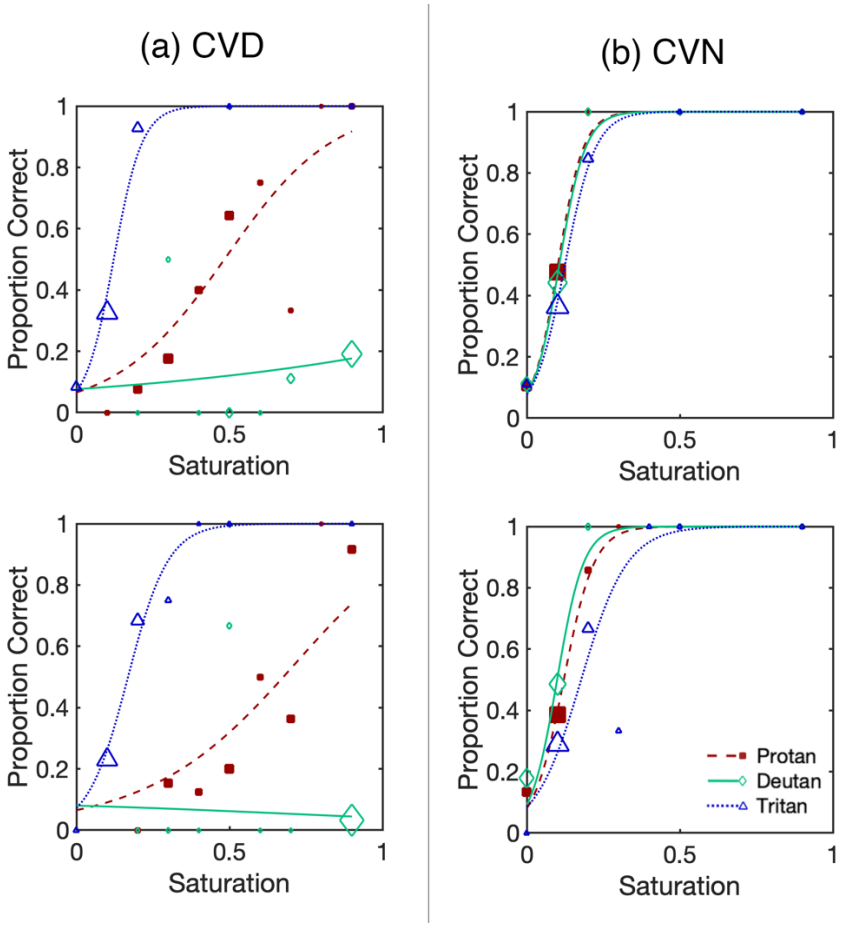


Figure 3.2. Examples of individual psychometric functions of four participants in the **(a)** CVD and **(b)** CVN groups. Each participant has three psychometric functions representing their performance for detecting protan (red squares, dashed line), deutan (green diamonds, solid line) and tritan (blue triangles, dotted line) targets as a function of *ColourSpot*'s stimulus saturation. The size of the data points is proportional to an arbitrary power of the number of trials of each saturation.

Comparing *ColourSpot's* CVD diagnosis with reported CVD diagnosis

Cohen's kappa (κ) was run to determine if there was agreement between *ColourSpot*'s diagnosis, parental reports and child self-reported diagnosis of CVD by an optician. There was almost perfect agreement between *ColourSpot* and parental reports of CVD diagnosis, κ=.792, 95% CI [0.61, 0.98], $p<.001$, and *ColourSpot* and child's self-reports of CVD diagnosis, κ=.841, 95% CI [0.67, 1.01], $p<.001$.

Comparison of CVD and CVN children on wellbeing and educational engagement

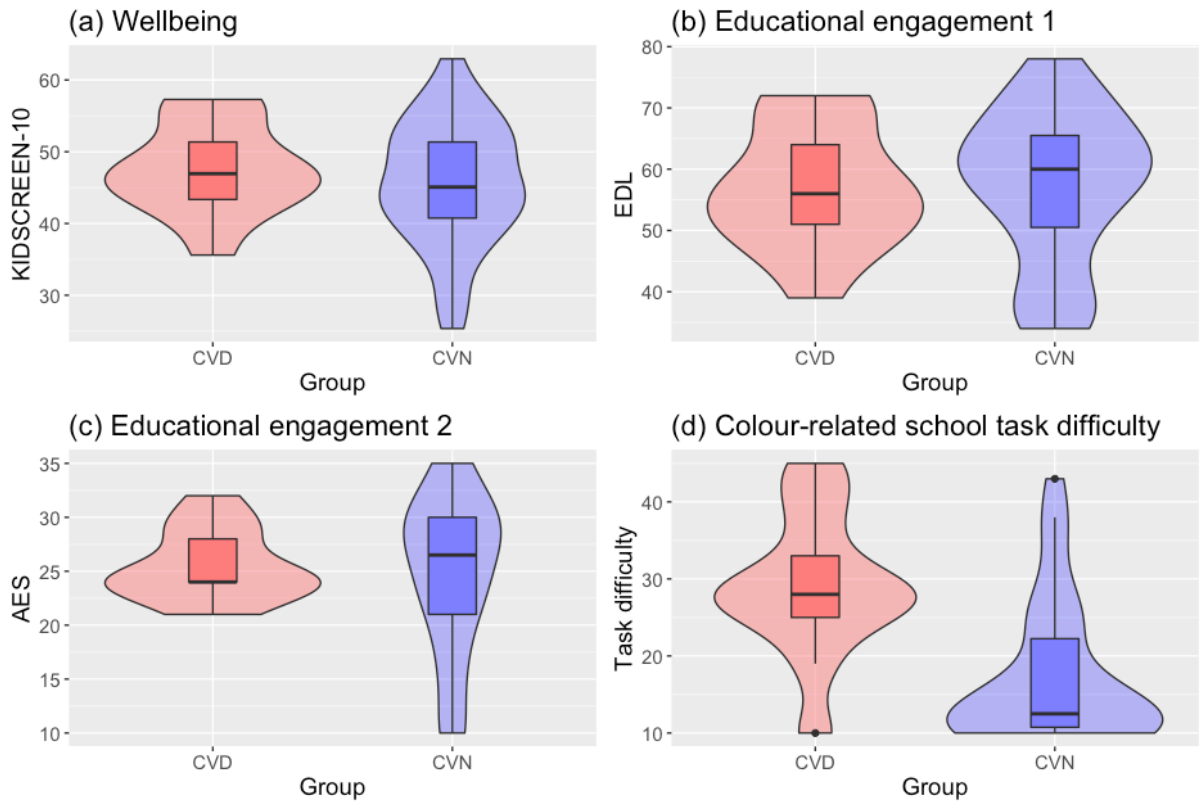


Figure 3.3. Violin boxplots showing the distribution, the median, and interquartile range of the data between the CVD and CVN group for **(a)** wellbeing (KIDSCREEN-10), **(b)** educational engagement 1 (EDL), **(c)** educational engagement 2 (AES) and **(d)** colour-related school task difficulty scores. The marker points for the colour-related task difficulty are outliers in the data.

Figure 3.3 shows violin boxplots of each questionnaire score, showing the distribution, median and interquartile range of the data between the CVD and CVN group. Using a non-parametric test to compare the differences between these two groups, a Mann-Whitney U test showed that KIDSCREEN-10 scores did not differ significantly from the CVD group (*Mdn*= 46.94) and the CVN group

(*Mdn*= 45.07), *U*= 178.50, *z*= -0.741, *p*=.0459, suggesting CVD does not have a negative impact on wellbeing. The EDL scores also did not significantly differ between CVD group (*Mdn*= 56.00) and the CVN group (*Mdn*= 60.00), *U*= 191.00, *z*= -0.426, *p*=.670. The AES scores also did not significantly differ between the CVD group (*Mdn*= 24.00) and the CVN group (*Mdn*= 26.50), *U*=200.00, *z*= -0.201, *p*=.841. Both suggesting CVD does not have a negative impact on educational engagement in the classroom. Further Bayesian hypothesis testing were used to assess the probability of the null hypothesis occurring compared to the alternative hypothesis to establish the strength of the data (Dienes, 2014; Van Den Bergh et al., 2020). Bayesian Mann-Whitney U tests using a Cauchy distribution centred around zero with a prior width distribution of 0.707 (Rouder et al., 2009) with 1000 iterations using Markov chain Monte Carlo (MCMC)'s sampling procedure, revealed anecdotal evidence for the null hypothesis, suggesting there is weak evidence in support of there being no effect between the CVD and CVN group for wellbeing (BF_{01}=2.67), educational engagement (BF_{01}=3.02) and agentic engagement (BF_{01}=3.10).

Comparison of CVD and CVN children for colour-related school task difficulty

There was a significant difference in colour-related difficulty task scores between the CVD group (*Mdn*= 28.00) and the CVN group (*Mdn*= 12.50), *U*=76.00, *z*= -3.32, *p*=.001. Thus, suggesting that older children and adolescents with CVD report significantly more difficulty with colour-related

tasks. Further Bayesian analyses revealed there was moderate evidence to support that there was a significant difference between CVD and CVN group on colour-related school task difficulties (BF_{10}=4.49). Figure 3.4 shows a bar graph comparing median scores for each question in the difficulties in colour-related school tasks in both groups. Multiple individual Bayesian Mann-Whitney U tests were conducted for each question to determine which colour-related task questions were significantly more difficult for the CVD group. Using a Cauchy prior distribution of 0.707 (Rouder et al., 2009) and based on 5 chains of 10000 iterations using MCMC's sampling procedure, Bayesian Mann-Whitney U tests revealed that colour-related school tasks like "Use colour in art and be confident in selecting the right colour" (BF_{10}=13.72) and "Name the colour of things" (BF_{10}=10.34) show strong evidence for the alternative hypothesis. Other tasks like "Do things in Science or Design Technology that require colour" (BF_{10}=9.05) and "Do Maths or English exercises when they are coloured (BF_{10}=4.66), show substantial evidence for the alternative hypothesis. And tasks like "Understand coloured diagrams, charts and figures" (BF_{10}=1.83) and 'Read letters and words when displayed in colour on the whiteboard or textbooks" (BF_{10}=2.33) show anecdotal evidence for the alternative hypothesis. See Figure 3.4.

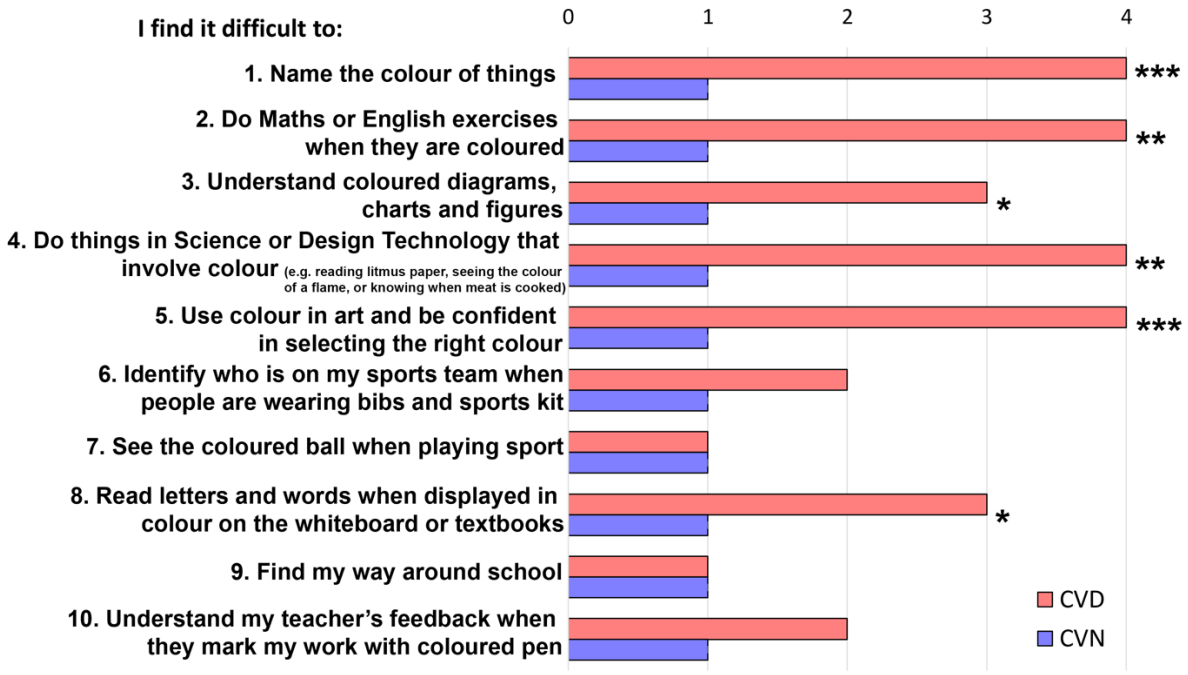


Figure 3.4. Comparing median scores between the CVD and CVN group for each question on the questionnaire for colour-related school task difficulty. Asterisks demonstrate questions that had Bayes factors in support of the alternative hypothesis including BF_{10} of 3 *, BF_{10} of 3-10 **, BF_{10} of 10 to 30 ***. Please note, a BF_{10} of 1 indicates no evidence; a BF_{10} of 1-3 indicates anecdotal evidence, a BF_{10} of 3-10 indicates substantial/moderate evidence, and a BF_{10} of 10-30 indicates strong evidence in support of the alternative hypothesis.

General ability to discriminate colour and its relationship to wellbeing and educational engagement

Irrespective of CVD, a continuous measure of general colour discrimination ability was calculated by taking the average of the protan, deutan and tritan thresholds. A high average colour discrimination threshold indicates poorer colour discrimination for protan, deutan and tritan targets. Figure 3.5 shows scatter graphs showing the correlational relationship between the average

colour discrimination thresholds and each questionnaire. A non-parametric Kendall's tau correlation was used to analyse the relationship between the average colour discrimination threshold and all questionnaire scores, as well as additional Bayesian analyses on Kendall's tau using a stretched beta prior width of 1.0 (Van Doorn et al., 2018). Analyses revealed there were no significant correlations between average colour discrimination with KIDSCREEN scores (τb=-0.03, p=.768, BF_{01}=4.95), educational engagement (EDL) (τb=-0.08, p=.422, BF_{01}=3.76), and agentic engagement (AES) (τb=-0.03, p=.80, BF_{01}=5.00), see Figure 3.5, with Bayes factors showing moderate evidence in support of the null hypothesis. However, there was a significantly positive correlation between average colour discrimination threshold and colour-related task difficulty scores, with extreme evidence supporting the alternative hypothesis (τb=0.38, p<.001, BF_{10}=141.37), suggesting poorer colour discrimination generally is associated with higher reports of colour-related school task difficulty.

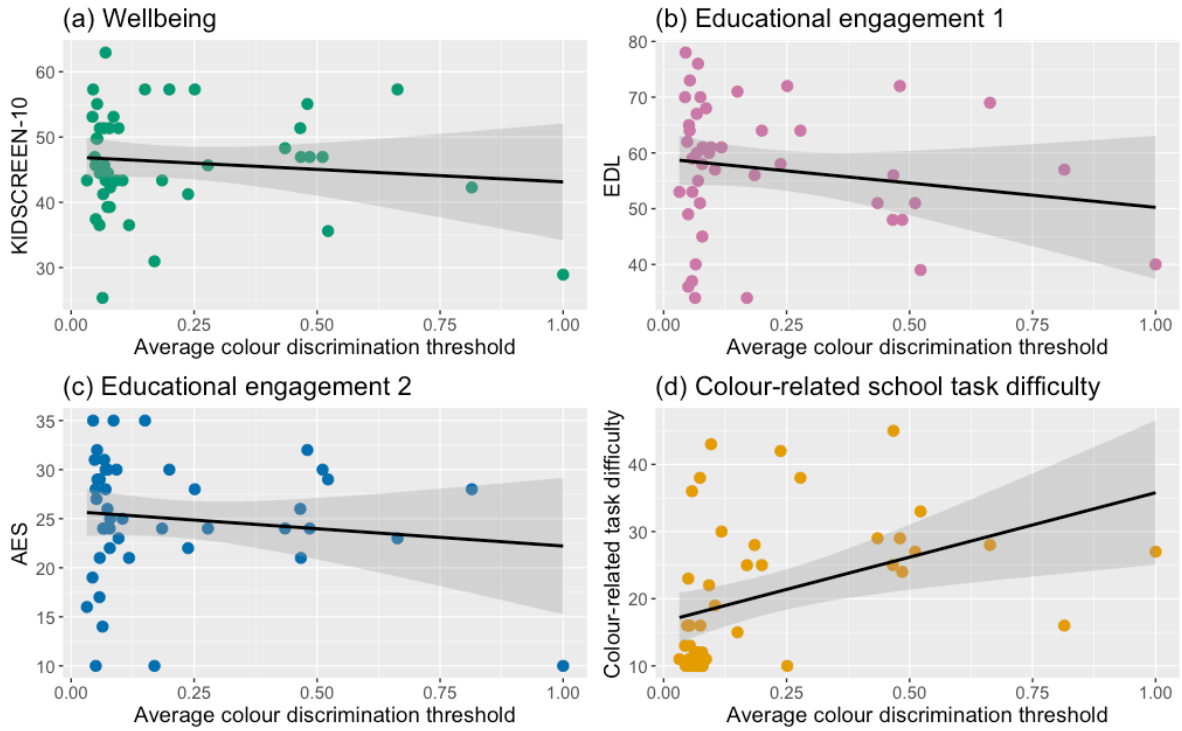


Figure 3.5. Scatter graphs showing the correlational relationship between average colour discrimination thresholds with **(a)** wellbeing (KIDSCREEN-10), **(b)** educational engagement 1 (EDL), **(c)** educational engagement 2 (AES) and **(d)** colour-related school task difficulty including the 95% confidence interval.

Does the extent of difficulty with colour-related school tasks correlate with wellbeing and educational engagement?

Further Kendall's tau correlations revealed there were no significant correlations between colour-related school task difficulty with wellbeing (KIDSCREEN-10), (τb=-0.156, p=.15, BF_{01}=1.70) or with educational engagement (EDL) (τb=-0.157, p=.142, BF_{01}=1.68), or with agentic engagement (AES) (τb=-0.176, p=.11, BF_{01}=1.25), see Figure 3.6. However, note that the Bayes factors reveal there is only anecdotal evidence in support of the null hypothesis, suggesting

that there is not sufficient evidence to make firm conclusions about the lack of a relationship.

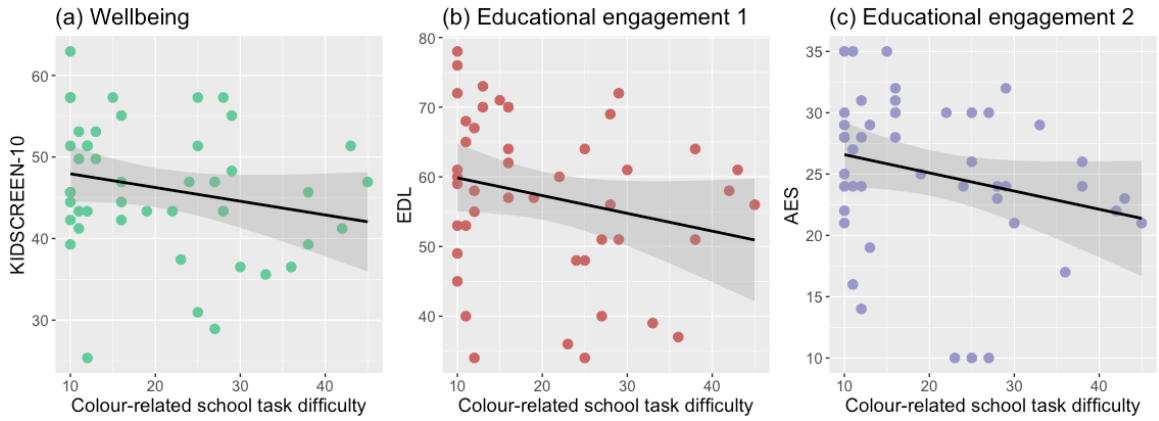


Figure 3.6. Scatterplots showing the relationship between colour-related school task difficulty with **(a)** wellbeing (KIDSCREEN-10), (**b)** educational engagement 1 (EDL) and **(c)** educational engagement 2 (AES) including the 95% confidence interval.

3.5 Discussion

This study aimed to establish whether CVD is associated with wellbeing and educational engagement issues in 11–16 year old boys and to quantify whether pupils with CVD self-report greater difficulty with colour-related school tasks than pupils who are CVN. The study provides a number of findings, although Bayesian analyses indicate that further evidence is required to make firm assertions about some of these findings. The first finding is that pupils with CVD and CVN score similar on wellbeing and educational engagement self-report questionnaires. However, Bayesian analyses suggest that the evidence for this null hypothesis is weak and that more data is required to make firm conclusions.

Second, the study finds that pupils with CVD report greater difficulties with colour-related school tasks compared to CVN pupils, and the Bayesian analyses suggest that there is moderate evidence for this claim. When considering specific tasks, there is strong evidence to suggest that children with CVD find it more difficult to “Use colour in art and be confident in selecting the right colour” and “Name the colour of things”. There is also substantial evidence to suggest that children with CVD find it more difficult to “Do things in Science or Design Technology that require colour” and “Do Maths or English exercises when they are coloured”. The use of *ColourSpot* also allowed us to quantify participants’ general colour discrimination ability irrespective of CVD. The third finding is that we found a positive correlation between average colour discrimination ability and difficulties in colour-related school tasks, with decisive evidence to support the alternative hypothesis, suggesting poorer colour discrimination is associated with higher self-reports of difficulty in colour-related school tasks. The final finding, which only had weak support from Bayesian analyses, was that pupils’ self-reported ability at doing colour-related tasks was not related to wellbeing and engagement in school. In sum, the current study fails to find any evidence that CVD or self-reported ability at doing colour-related tasks affects wellbeing and educational engagement in 11–16 year old boys, although Bayesian analyses suggests that further data collection is needed before firm statements about a lack of effect can be made. In contrast, the study does find firm evidence that 11–16 year old boys with CVD self-report greater difficulty at doing colour-related tasks in school, in particular those that involve naming colours and using colour in art.

The current study fails to find evidence that CVD is associated with lower wellbeing, which is in contrast to several previous studies (Thomas et al., 2018; Thuline, 1964) but in support of others (Nithiyaananthan et al., 2020). The Bayesian analysis indicates that there is insufficient data in the current study to be confident in the lack of effect, and a replication of the current study with a larger sample would be sensible. If the lack of effect proves to be robust, one question is why there is discrepancy with some other studies. One possibility is that the effect of CVD on wellbeing, behaviour, educational achievement and engagement varies according to cohort (e.g., age or country of the participants), or that some questionnaires are more sensitive at detecting differences than others. Whether or not the questionnaire relies on self-report, parent-report or teacher-report may also influence the finding. For example, one study which found an effect of CVD on behavioural and emotional problems used the CBCL which relies on parent-report (Thomas et al., 2018), yet another study did not find behavioural and emotional differences using the teacher and self-reported SDQ (Nithiyaananthan et al., 2020).

One important factor which has not been considered in research so far is whether the child's awareness of their CVD affects the impact of their CVD on their wellbeing, educational achievement or quality of life. In the current study, all the pupils and their parents (except one participant) were already aware that they have CVD. It is possible that this awareness may mitigate the negative

impact of CVD on an individual's wellbeing and educational engagement. For example, if a child has been diagnosed with CVD, teachers and parents may provide support for this by accommodating the CVD with different teaching materials or simple measures such as labelling colouring pencils. A child's difficulty in the classroom with colour tasks may also be better understood if the child has a CVD diagnosis, causing less frustration and therefore less negative impact on wellbeing or educational engagement. Whether or not CVD diagnosis is protective of the negative impact of CVD on wellbeing and educational engagement could be tested directly in further studies by having separate groups of CVD participants who are and who are not aware of their CVD status. Another question which has not been directly investigated is whether the potential for a negative impact of CVD on wellbeing or quality of life accumulates as an individual gets older, whether a particular stage in life is more affected, or whether individuals learn protective strategies as they develop. Further studies that take a developmental approach would be useful to address this question. The one finding that the current study can be decisive on is that 11–16 year old boys with CVD report significantly more difficulty with colour-related school tasks than CVN children. This finding supports other studies that highlight that CVD pupils report that they struggle with particular colour-related school tasks and quantifies further reports from individuals with CVD (Albany-Ward, 2011; Chan et al., 2014; Mashige, 2019; Steward & Cole, 1989; Sullivan, 2011; Ugalahi, Fasina, & Ogun, 2016; Waggoner, 2017). In the current study, the CVD group reported particular difficulty with art and naming colours, as well as using colour in science or design technology, and completing

maths and English exercises when coloured, but there were no differences in other tasks such as identifying teams in sports as has been found in previous studies (Mashige, 2019; Ugalahi, Fasina, & Ogun, 2016). These studies are potentially useful in highlighting that children with CVD may lack confidence in certain tasks at school and need additional support. However, again it is unclear whether this lack of confidence comes from the participants being aware of their CVD and therefore expecting to be poorer on these tasks, or whether children with CVD have an accurate self-assessment of their abilities. A study which objectively measures the ability of CVD pupils to complete various school tasks would clarify this and would quantify the extent of any difficulty experienced.

Another important question for further research is to consider the importance of the severity and type of CVD on the impact that it has in the school environment. For example, dichromats have poorer colour vision than anomalous trichromats, and therefore we expect their CVD to have a greater impact on an individual's life. We also know that the impact of anomalous trichromacy on colour vision can be compensated for by cortical amplification of the chromatic signal at the cortex and therefore the resulting chromatic experience may be less altered than one might expect on the basis of the functioning of the cones (Tregillus et al., 2021). The current study's use of *ColourSpot* enabled us to quantify an individual's colour ability on a continuous measure, yet a participant's CVD was not classified beyond that. A future study

using the gold standard anomaloscope for diagnosis would enable classification of CVD and the impact of different types of CVD could be compared.

Further research on the impact of CVD on the quality of life of school pupils would be worthwhile as it would enable a well-informed decision on whether colour vision screening should be implemented in schools. In the UK, the decision to stop screening for CVD in schools was based on insufficient evidence (Cumberland et al., 2004), none which are relevant to the modern colour coded classroom of today. Future research should address the question of whether diagnosis of CVD at an early age in school helps mitigate the negative effects of CVD on quality of life that have been revealed in adulthood (Barry et al., 2017; B. Cole, 2015; Long et al., 2015; Steward & Cole, 1989). Although the current study failed to find evidence that CVD affects wellbeing or educational engagement, there was firm evidence that children with CVD feel that they struggle more with certain colour-related school tasks. Further research into the types of tasks that children with CVD struggle with or lack confidence in could help develop a child educational version of the quality of life impact of colour blindness (CBQoL) questionnaire that was developed for assessing the impact of CVD in adults (Barry et al., 2017). Such a tool would be useful for educators to understand the challenges that pupils with CVD face and help develop education plans on how to support them. For now, a handful of previous studies, anecdotal evidence and some of the findings in the current

study suggest that the impact of CVD on school children is an issue worth taking seriously.

Chapter 4

Paper 3: The contribution of various pattern characteristics and image statistics to visual preference in infants and toddlers

Teresa Tang[1], Alice Skelton[1], Jenny Bosten[2] & Anna Franklin[1]

[1] The Sussex Colour Group, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

[2] Sussex Vision Lab, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

4.1 Abstract

Visual pattern preferences displayed early in development have provided insight into the development of the visual system and the perceptual capabilities of infants and young children (Fantz, 1963, 1965; Fantz & Miranda, 1975). The current study further investigates visual pattern preference in infants and toddlers and asks whether certain pattern characteristics such as odd-one-out features, and chromatic and spatial image statistics affected the amount of looking in the two age groups. Eye-movements of 6-8 month old infants (N=22) and 2-3 year old toddlers (N=20) were recorded whilst they were presented with a set of patterns. Toddlers looked significantly longer than infants overall, but for a similar time than infants at patterns with faces and odd-one-out features and longer than infants at patterns with animals, trees and abstract shapes. The dominant colour in a pattern contributed to visual preference in infants and toddlers, but other chromatic and spatial image statistics did not predict looking

at either age. The findings have implications for our understanding of the development of visual pattern preference and can be applied to the design of optimal patterns for infants and young children. The study also provides an analysis template for further research on the role of image statistics in visual development.

Keywords: visual development, image statistics, image analysis, developmental visual aesthetics, pattern preference

4.2 Introduction

In the first months of life, infants' vision rapidly develops. From birth, visual acuity improves at least five-fold by 6 months old (Courage & Adams, 1990; Maurer et al., 1999; Mayer & Dobson, 1982; van Hof-van Duin & Mohn, 1986) and contrast sensitivity increases approximately 10-fold by 8-9 months old (Atkinson et al., 1977; Dekker et al., 2020; Norcia et al., 1990; Peterzell et al., 1995). Infant colour vision also rapidly develops in the first few months of life and infants are fully trichromatic by at least 3 months old (Catherwood et al., 1990a; Knoblauch et al., 2001; Morrone et al., 1990; Suttle et al., 2002; Teller & Bornstein, 1987). Despite initially having poor vision, infants can discriminate patterns (Fantz, 1961, 1965; Fantz & Miranda, 1975). For example, faces are one of the most complex visual stimuli, yet neonates and infants show an early or innate-like biased preference for faces and face-like schematic stimuli

(Cassia et al., 2004; Dannemiller & Stephens, 1988; Farroni et al., 2005; Goren et al., 1975; Johnson et al., 1991; Simion et al., 2001, 2007; Simion & Giorgio, 2015; Sugita, 2009). Studies have consistently demonstrated that infants have a strong preference to faces compared to other stimuli and this strong face bias preference continues from infancy to adulthood (Dannemiller & Stephens, 1988; Frank et al., 2009, 2014; Gliga et al., 2009; Johnson, 2005; Nelson, 2001; Scherf & Scott, 2012; Simion & Giorgio, 2015).

In addition to faces, infants show other distinctive pattern preferences which change throughout development. The classic experimental works of Robert Fantz in the 1950s to 1970s were some of the first to demonstrate infant pattern preferences, and this demonstration of reliable pattern preferences at a very young age was critical to changing the opinion that infant's visual world was a "blooming, buzzing confusion" (James, 1890). For example, Fantz (1963) found that neonates preferred black and white patterns such as schematic faces, concentric circles and newspaper print compared to plain coloured stimuli. It was also demonstrated that after 2 months old, infants prefer bulls-eye patterns compared to horizontal striped patterns, and prefer looking at checkerboard patterns compared to shapes like triangles, crosses, squares and circles (Fantz, 1958, 1961; Spears, 1964). Other visual preferences have since been revealed. For example, infants prefer patterns with high contrast (Banks & Dannemiller, 1987; Stephens & Banks, 1987), and prefer curved shapes and circular patterns compared to straight rectangular edges (Amir et al., 2011; Fantz & Miranda,

1975; Fantz & Nevis, 1967; Jadva et al., 2010). Visual stimuli with odd-one-out features will have a pop-out effect and capture infant's attention after 3 months old (Adler et al., 1998; Adler & Orprecio, 2006; Catherwood et al., 1996; Colombo et al., 1995; Gerhardstein et al., 1999; Quinn & Bhatt, 1998; Rieth & Sireteanu, 1994; Rovee-Collier et al., 1992; Salapatek, 1975). Infants aged 4-6 months old also have specific colour preferences, looking longest at blue and least at yellow-green (chartreuse) when presented individually or in pairs (Bornstein, 1975; Brown & Lindsey, 2013; Franklin et al., 2008, 2010; Skelton & Franklin, 2020; Teller et al., 2004; Zemach & Teller, 2007). By 12 months, infants prefer patterns with vertical symmetry compared to horizontally symmetrical and asymmetrical patterns (Bornstein et al., 1981; Bornstein & Krinsky, 1985; Fisher et al., 1981).

From 3 weeks to 3 months old, infants develop an increasing preference for complex patterns (Brennan et al., 1966; Greenberg & Blue, 1975). There also appears to be an optimal complexity preference, where infants and children like to allocate their attention to visual stimuli with intermediate complexity that is neither too simple or too complex (Cubit et al., 2021; Kidd et al., 2012; Poli et al., 2020). Although the relationship between patterns and visual art preferences in infancy is largely unclear (Göksun et al., 2014), the preference for intermediate complexity can be seen in aesthetic preferences for paintings, where infants prefer looking at original art compared to when the original art's complexity was manipulated (Krentz & Earl, 2013). Infants also prefer paintings

by Picasso compared to Monet (Cacchione et al., 2011) and when 11-12 year old children viewed paintings by van Gogh, children tended to rely more on bottom-up processing (using low-level features to guide attention) compared to adults who were more likely to rely on top-down processing (using prior knowledge to guide attention) (Walker et al., 2017). This developmental change from more bottom-up to top-down processing is also seen in children and adults when viewing natural images (Açık et al., 2010). Children also show an increasing gradual preference for natural environment images compared to urban environment images from 4-11 years old (Meidenbauer et al., 2019). Hence, infants both show early distinctive pattern preferences and preferences for complex images which change during development. The development of these visual preferences has potential to provide insight into the principles that guide visual and perceptual development.

One underexplored question on visual development is the extent to which the developing visual system tunes into statistical regularities of low-level and high-level properties of scenes. Statistical regularities in the properties of scenes and images, such as hue, symmetry and luminance, are termed “image statistics”. Existing infant research has mostly investigated whether saliency, a high-level image statistic composed of low-level image statistics of colour, orientation, and motion (Itti & Koch, 2001; Walther & Koch, 2006), can predict infant looking (Amso et al., 2014; Franchak, 2020; Franchak et al., 2016; Frank et al., 2009; Kadooka & Franchak, 2020; Kwon et al., 2016). For example, studies have

found that due to bottom-up processing, infant looking behaviour while watching children's television is predicted by visual salience (Franchak et al., 2016; Frank et al., 2009). However, the contribution of various low-level image statistics to infant visual preferences is unclear.

Adult vision is known to be optimised to the statistical regularities of natural scenes which contain natural image statistics (Olshausen & Field, 2000; Simoncelli & Olshausen, 2001). The field of visual aesthetics has shown that natural image statistics affect preferences not only for natural scenes (Berman et al., 2014; Ibarra et al., 2017; Kardan et al., 2015), but also preferences for art and paintings (Graham & Field, 2007; Graham & Redies, 2010; Mallon et al., 2014; Montagner et al., 2016; Nascimento et al., 2021). The hypothesis is that if the visual system is optimised to process and adapt to the statistical properties of the natural world, then art or abstract images with natural visual statistical properties will also be processed by the visual system more efficiently and hence be highly preferred (Graham & Redies, 2010; Redies, 2007; Simoncelli & Olshausen, 2001). There is some support for this, for example, natural images have a distinct chromatic distribution that is aligned with the blue-yellow daylight axis (Bosten et al., 2015; Burton & Moorhead, 1987; McDermott & Webster, 2012; Ruderman et al., 1998; Shevell & Kingdom, 2008), and adults rate abstract chromatic Mondrian patterns with the same natural blue-yellow colour distribution as most aesthetically pleasing than patterns of other colour distributions (Juricevic et al., 2010). Similarly, if natural images are decomposed

to their spatial components using a Fourier transform, the amplitude spectral slope of natural images show a distinct $1/f^{\alpha}$ spatial frequency (or equivalently $f^{-\alpha}$), where α (the amplitude spectral slope) is approximately 1 (Burton & Moorhead, 1987; D. Field & Brady, 1997; Hansen & Hess, 2006; Tolhurst et al., 1992; Van Hateren & Ruderman, 1998); and art or abstract images with this natural spatial characteristic are preferred by adults (Isherwood et al., 2021; Juricevic et al., 2010; Nguyen & Spehar, 2021; O'Hare & Hibbard, 2011). However, the developmental trajectory for sensitivity to natural image statistics is unclear. There is some evidence that sensitivity to natural chromatic statistics (Skelton et al., 2021) and natural texture statistics is present in infancy and childhood (Balas et al., 2018, 2020), but the sensitivity to natural fractal patterns is present at 3 years old (Robles et al., 2020), and sensitivity to the natural amplitude spectral slope ($1/f$) does not fully develop until approximately 10 years old (Ellemberg et al., 2012). One potential way of further understanding the role of natural image statistics in visual development is to explore whether image statistics predict infant looking preferences for art, abstract images or patterns, as has been demonstrated for adult preference. One possibility is that natural image statistics could explain the distinct visual pattern preferences seen in infancy and early childhood.

The current study further investigates visual pattern preferences in infants and young children to understand how visual characteristics of the patterns and their image statistics predict looking in two developmental age groups. Eye-

movements were recorded whilst 6-8 month old infants and 2-3 year old toddlers viewed a set of 13 patterns, designed by the baby product company Cosatto Ltd. These patterns were chosen since they vary in colour and shape and have a range of characteristics such as odd-one-out features, schematic faces, animals, and abstract shapes. Based on previous literature (Adler & Orprecio, 2006; Frank et al., 2009; Gliga et al., 2009; Quinn & Bhatt, 1998; Simion & Giorgio, 2015), we predict that infants will look longer at patterns with odd-one-out features and schematic faces than other patterns. We also investigate whether low-level and high-level image statistics of the patterns predict how long infants and toddlers look, and whether there is a bias for longer looking at patterns that have natural image statistics.

4.3 Methods

Participants

Thirty 6–8 month old infants (17 females, *M*= 7.33 months, *SD*= 0.43) and 22 2–3 year old toddlers (9 females, *M*= 2.91 years, *SD*= 0.53) participated in the study with 10 children (8 infants, 2 toddlers) excluded in the final sample due to insufficient number of trials completed. The final sample included 22 infants (13 females, *M*= 7.24 month, *SD*= 0.46) and 20 toddlers (8 females, *M*= 2.94 years, *SD*= 0.51). Parents of infants and toddlers were recruited online via Sussex Baby Lab's social media. All infants were born full-term, had a minimum birth weight of 2500g and had no known visual disorders.

Stimuli

Thirteen patterns produced by Cosatto Ltd, a British brand which designs prams, pushchairs and other baby products, were presented to infants and toddlers (Figure 4.1). The patterns are designs used in the liners of Cosatto's car seats, prams and pushchairs products. Patterns were displayed on a 22-inch Diamond Pro 2070SB CRT monitor (Mitsubishi, Tokyo Japan, (screen resolution 1600 x 1200 pixels at 85 Hz refresh rate)). Eye movements were recorded with an EyeLink 1000 Plus eye-tracker (SR Research, Ontario, Canada) and the experiment was run in Experiment Builder version 2.1.140 (SR Research, Ontario, Canada).

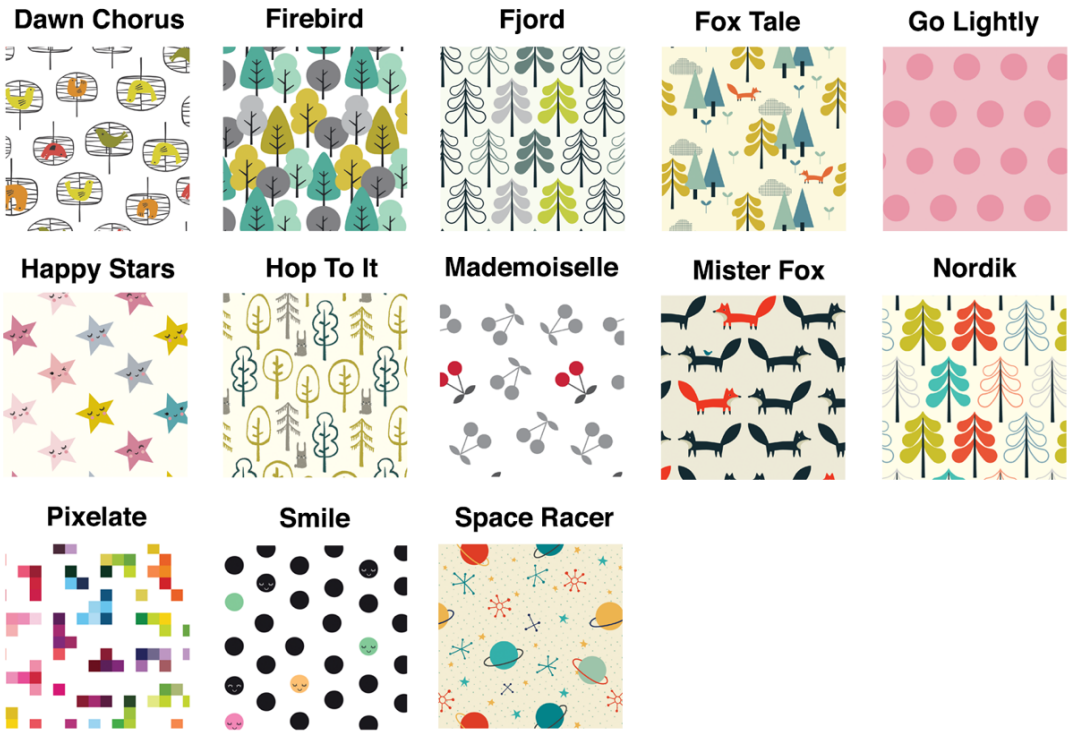


Figure 4.1. All 13 patterns by Cosatto Ltd with their design name.

Design and procedure

Infants and toddlers sat on at a car seat secured on a chair approximately 60cm away from the monitor. Looking was calibrated using a 4-point calibration procedure where a looming achromatic motion target was presented at four locations of the screen. Each pattern image was presented once for 5000 milliseconds (ms) presented with a visual angle of 10° (screen dimensions= 495 x 484.5 mm, eye to screen distance= 600mm) against a grey background. Infants and toddlers had to complete all 13 trials to be included in the final sample. A looming achromatic motion target was used as a visual attention grabber and was presented after each trial to maintain the participant's gaze at the centre of the screen. Families were gifted a small Cosatto toy for infants and a children's book for toddlers as compensation for their time and contribution.

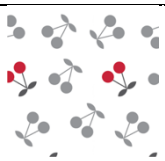
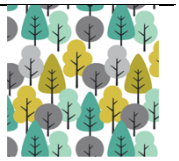
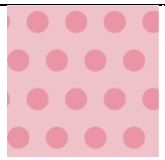
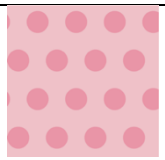
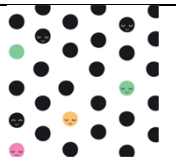
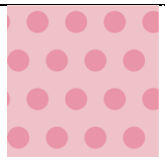
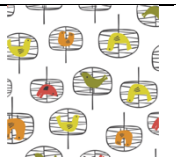
Image processing and image statistics

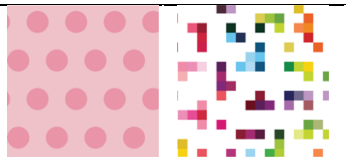
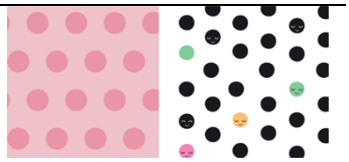
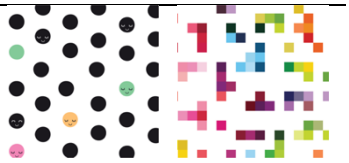
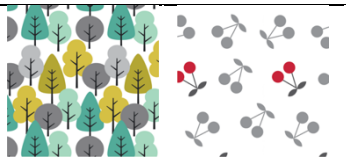
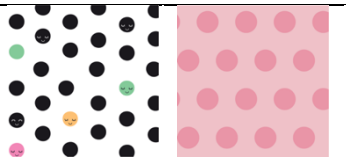
Image files were in Portable Network Graphic (.PNG) format and 1182 x 1182 in pixel size. Image statistics were calculated (see Table 4.1) and processed via MATLAB's Image Processing Toolbox. Spatial image statistics were selected based from previous literature (Bornstein & Krinsky, 1985; Ibarra et al., 2017; Kardan et al., 2015; Mather, 2018), including spectral slope, entropy, edge density and symmetry. Feature congestion was also included as a measure of visual clutter, a high-level image statistic similar to saliency (Rosenholtz et al., 2007; Wass & Smith, 2015). Chromatic image statistics were converted from

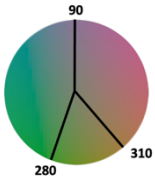
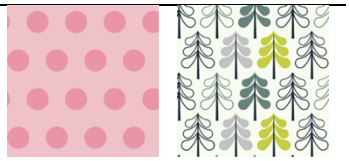
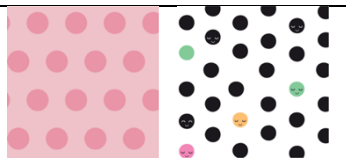
RGB matrices to LMS matrices in MacLeod-Boynton colour space to be appropriate for human colour vision perception. LMS matrices were based on Stockman, MacLeod, and Johnson's 10-degree cone fundamentals (Stockman et al., 1993), which are predicted spectral sensitivities of L, M and S cones. Images were also gamma corrected to match the display monitor. Table 4.1 provides a full list of the image statistics, an explanation of these and example patterns with low and high values for each image statistic.

Table 4.1. Definition of image statistics, an explanation of how they are calculated, and an example of a pattern with a low and high image statistic of each image statistic group. All were computed via MATLAB, using the Image Processing Toolbox and other built-in MATLAB functions.

Image statistics	**Definition**	**Example patterns with low or high image statistic**	
Spatial		**Low**	**High**
Spectral slope	A Fourier transformation is computed on a greyscale image (L+M matrix) and the spectral slope of the fitted line to the rotationally averaged amplitude spectrum on log-log coordinates is extracted (Mather, 2018, 2020). Natural images usually have an amplitude spectral slope of approximately -1 (Burton & Moorhead, 1987; D. Field & Brady, 1997; Hansen & Hess, 2006; Tolhurst et al., 1992; Van Hateren & Ruderman, 1998). A steeper slope means the image has a greater amount of low spatial frequency compared to high spatial frequency.		

Entropy	A measure of the randomness of an image, measuring the extent the information in an image varies unpredictably. Entropy is calculated from the intensity of a grayscale histogram of the image, where each pixel could have an intensity value of 0 to 255 (8-bit-grayscale; Berman et al., 2014). The more the intensity histogram shows that all intensity values occur with the same probability in the image, the higher the entropy value of an image. Natural environment images tend to have higher entropy than urban environment images (Berman et al., 2014; Kardan et al., 2015).		
Edge density	The edge density ratio calculates the ratio of edge pixels to the total number of pixels in an image. Hence, the higher the edge density ratio the higher number of edges in the image. This is calculated by using MATLAB's edge function using the Canny method (Canny, 1986) which filters noise and detects strong and weak edges (Berman et al., 2014).		
Symmetry	Flips the image left to right (vertically) and compares how similar each image is in a pixel-by-pixel comparison by calculating the mean squared error. Hence, the smaller the number the more vertically symmetrical the image is.		
Visual clutter			
Feature congestion	An established measure of visual clutter in an image, using colour, contrast and orientation (Rosenholtz et al., 2007). The local clutter in colour (using CIELab, a perceptual colour space (CIE, 1977)), contrast and orientation are computed and combined and then combined over space using a Minkowski mean of order and then averaged to output a single measure of feature congestion. A		

	higher number means an image has greater feature congestion.	
Chromatic		
SD hue	Demonstrates the variation of hue using the standard deviation of the hues of every pixel in an image using circular statistics (Philipp, 2009). Hue is the perceptual property of colour corresponding to the physical property of the dominant wavelength of colour (Burns & Shepp, 1988).	
SD luminance	Demonstrates the luminance variance in an image, using the standard deviation of the total luminance value (L+M matrix) of each pixel. Luminance is the amount of brightness in a colour. Hence, a higher number indicates more variation in brightness.	
SD saturation	The standard deviation of the saturation of all the pixels in the image. Saturation is the intensity of colour. A higher number means the larger variation of saturation in an image.	
Mean saturation	The average saturation of all the pixels in the image. Calculated based from the specified white point in the MacLeod-Boynton chromaticity diagram. A higher mean value means a larger average saturation in an image.	
Mean luminance	The average total luminance value of each image pixel, using the L+M matrix of the image. A higher mean value indicates a greater average luminance in an image.	
Log axis ratio	Measures the direction of colour distribution in an image, regardless of the colour distribution's location or the total area of the colour distribution. A log axis ratio > 0 indicates more colour variation around the blue-yellow colour axis, which is typically seen in natural images (Burton & Moorhead, 1987; McDermott & Webster, 2012;	

	Ruderman et al., 1998). A log axis ratio < 0 indicates more colour variation around the orthogonal red-green colour axis. A log axis ratio equal to 0 indicates equal variation in both colour directions. Hence, 'Go Lightly' has a large log axis ratio even though the pattern is mostly pink as the colour distribution of the image mainly varies along the blue-yellow colour axis (see SI, Figure S7) and the colours in the image do not have a large varied colour distribution.	
Hue bins	A hue bin categorises the colours in the image using hue angles specified in polar MacLeod-Boynton colour space. It calculates the percentage of pixels in the image which fall within a specified hue angle. We calculated three hue angles (see below) based on the colour distribution of all the patterns (see SI, Figure S7).	
Pink-reddish hues	The percentage of pixels of an image that are reddish (violet-magenta-red-orange), specified in hue angle 90-310° in MacLeod-Boynton colour space in polar coordinates.	
Chartreuse-ish hues	The percentage of pixels of an image that are chartreuse-ish (orange-yellow-green), specified in hue angles 310-280° in MacLeod-Boynton colour space in polar coordinates.	
Cyan-blue-ish hues	The percentage of pixels of an image that are cyan-bluish (green-cyan-blue), specified in hue angles 90-280° in MacLeod-Boynton colour space in polar coordinates.	
Achromatic hues	Achromatic colours (i.e., black, white, and grey) in the images that are not specified in the hue bins. Calculated from the total sum of all specified hue bins, subtracted from 1 (1- total of pixels in hue bins).	

Statistical analysis

Eye-movement data were extracted from Eyelink Data Viewer (SR Research, Ontario, Canada), where looking time was measured in dwell time (the total amount of time spent looking at the pattern), using the total sum of all fixation durations (ms)). The total fixation counts for each trial were also calculated but later excluded as their pattern of results was similar to dwell time. Firstly, a mixed analysis of variance (ANOVA) was used to see whether looking time varied across the patterns for infants and toddlers. Further post-hoc pairwise comparisons were then analysed to compare the looking time differences across patterns between infants and toddlers. Secondly, all the patterns were categorised into five pattern types and a mixed ANOVA analysed the effect of pattern type on looking time between infants and toddlers. Further post-hoc pairwise comparisons were made to compare looking differences for image types between infants and toddlers and independent post-hoc one-way ANOVAs were used to analyse the pattern of looking between images in each age group. A Greenhouse-Geisser correction was applied for all ANOVA analyses if Mauchly's test of sphericity was violated (p<.05). The image statistics were grouped into four multiple linear regression models to analyse whether image statistics could predict looking time in infants and toddlers, measuring: (1) spatial image statistics (spectral slope, edge density, entropy and symmetry), (2) feature congestion, (3) colour variance (SD hue, SD luminance, SD saturation) and (4) colour properties (mean saturation, mean luminance and log axis ratio). To further understand the effects of hue, the most dominant hue was identified in each image and each image was categorised

into one of four chromatic hue types (pink-reddish, chartreuse-ish, cyan-blue-ish and achromatic hues), enabling us to use a mixed ANOVA to analyse whether the most dominant hue type in an image affected looking time in infants and toddlers. Using a multivariate fixed effects prior distribution (r scale prior width of 0.5) for Bayesian ANOVA tests (Rouder et al., 2012) and a Cauchy prior distribution of 0.707 for Bayesian t-tests (Rouder et al., 2009), Bayes factors were also reported throughout analyses to determine the strength of the data by assessing the probability of the null hypothesis occurring compared to the alternative hypothesis (Dienes, 2014; Van Den Bergh et al., 2020).

4.4 Results

Does looking time vary across the patterns for infants and toddlers, and do infants and toddlers vary in their pattern preferences?

Figure 4.2 gives the mean dwell time for infants and toddlers for each pattern.

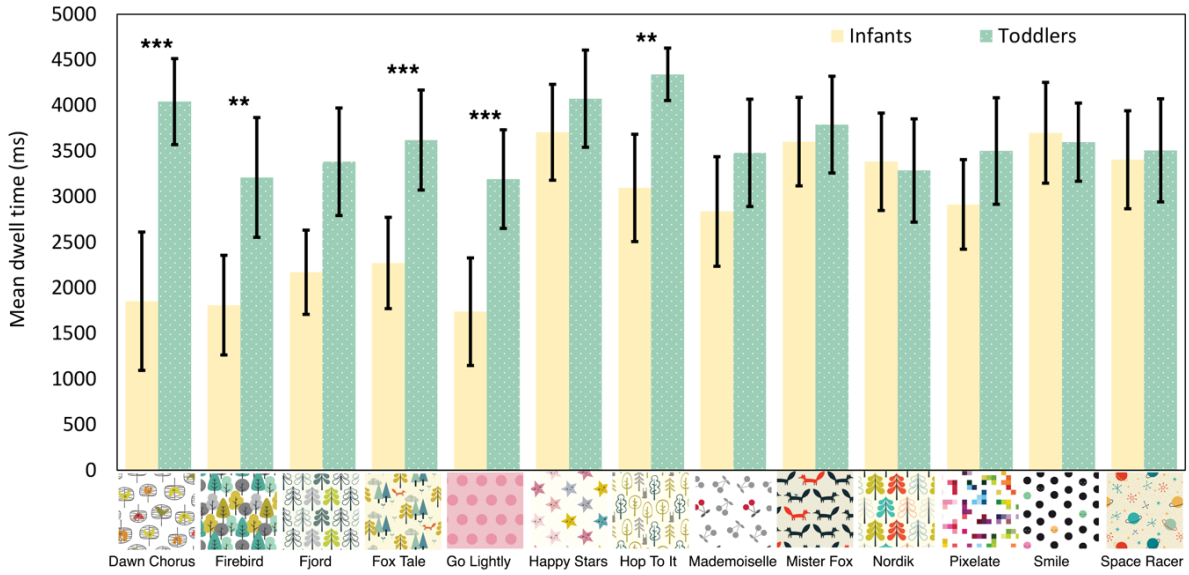


Figure 4.2. The mean dwell time (ms) infants and toddlers spent looking at each pattern, including the 95% confidence interval (CI) error bars. Asterisk (*) indicates significant difference between infant and toddler dwell time to patterns (** p<.01, ***p <.001).

To measure whether there were any significant differences in looking across the patterns for infants and toddlers, and to see whether the looking pattern varies with age, a 13 X 2 mixed ANOVA analysed the effect of patterns (within-subjects factor) and age (between-subjects factor) on dwell time. Results revealed that dwell time differed across the patterns, with a significant main effect of patterns on dwell time, F (12,480) =7.13, p<.001, η^2=0.10, BF_{10}= 808100000. There was also a significant main effect of age on dwell time, F (1,40) =14.65, p<.001, η^2=0.08, BF_{10}=44.97, where toddlers significantly spent more time looking at patterns than infants (p<.001, BF_{10}=378200000). There was also a significant interaction of pattern and age, F (12,480) =4.195, p<.001, η^2=0.10, BF_{10}=7595.20. Post-hoc Bonferroni corrected paired samples t-tests

revealed that toddlers significantly spent more time than infants looking at 'Dawn Chorus' (p<.001, BF_{10}=330.11), 'Fjord' (p=.002, BF_{10}=20.58), 'Fox Tale' (p<.001, BF_{10}=78.46), 'Go Lightly' (p<.001, BF_{10}=42.26), and 'Hop To It' (p=.002, BF_{10}=25.16), see Figure 4.2.

These interactions were further explored with post-hoc one-way repeated ANOVAs to confirm whether there were significant differences in looking across patterns independently within each age group. One-way repeated measures ANOVAs confirmed that infants significantly varied in their dwell time across patterns, with a significant main effect of patterns on dwell time, F (12, 252) = 9.07, p<.001, η^2=0.30, BF_{10}=604500000000. Toddlers also significantly varied in their dwell time across patterns, F (12,228) =2.13, p=.016, η^2=0.10, although Bayes factor indicated anecdotal evidence in support of the alternative hypothesis (BF_{10}=1.56).

Do infants and toddlers differ in their looking preferences for different pattern types?

Patterns were classified as one of five types: patterns that have: (1) animals, (2) faces, (3) odd-one-out features, (4) trees and (5) abstract shapes (see Figure 4.3). An average dwell time was calculated for each pattern type for all participants in both age groups.

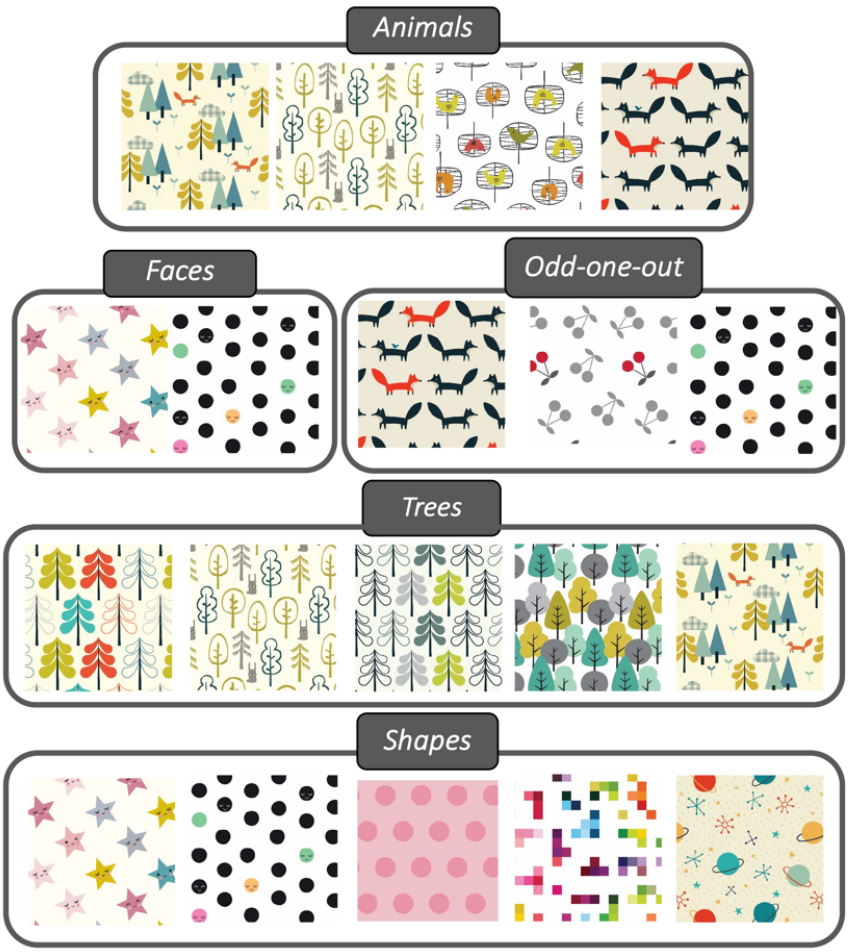


FIGURE 4.3. Five pattern types: Patterns with animals (four patterns): 'Fox Tale', 'Hop To It', 'Dawn Chorus', 'Mister Fox'; patterns with faces (two patterns): 'Happy Stars', 'Smile'; patterns with odd-one-out features (three patterns): 'Mister Fox', 'Mademoiselle', 'Smile'; patterns with trees (five patterns): 'Nordik', 'Hop To It', 'Fjord', 'Firebird', 'Fox Tale'; and patterns with abstract shapes (five patterns): 'Happy Stars', 'Smile', 'Go Lightly', 'Pixelate', 'Space Racer'.

To analyse whether infant and toddler looking preferences differed between pattern type, two 5 x 2 mixed ANOVAs analysed the effect of pattern type (within-subjects levels: faces, animals, trees, odd-one-out, abstract shapes) and age (between-subjects levels: infants, toddlers) on average dwell time.

Regardless of age, the average dwell time significantly differed across the five pattern types, F (3.0, 119.84) =13.76, p<.001, η^2=0.07, BF_{10}=523487.04. Further post-hoc Bonferroni corrected pairwise comparisons revealed significantly more time was spent looking at patterns with faces compared to patterns with animals (p<.001, BF_{10}=33.47), trees (p<.001, BF_{10}=13838.72) and abstract shapes (p<.001, BF_{10}=4024.77), and that significantly more time was spent looking at patterns with odd-one-out features compared to patterns with trees (p<.001, BF_{10}=13.16), see Figure 4.4.

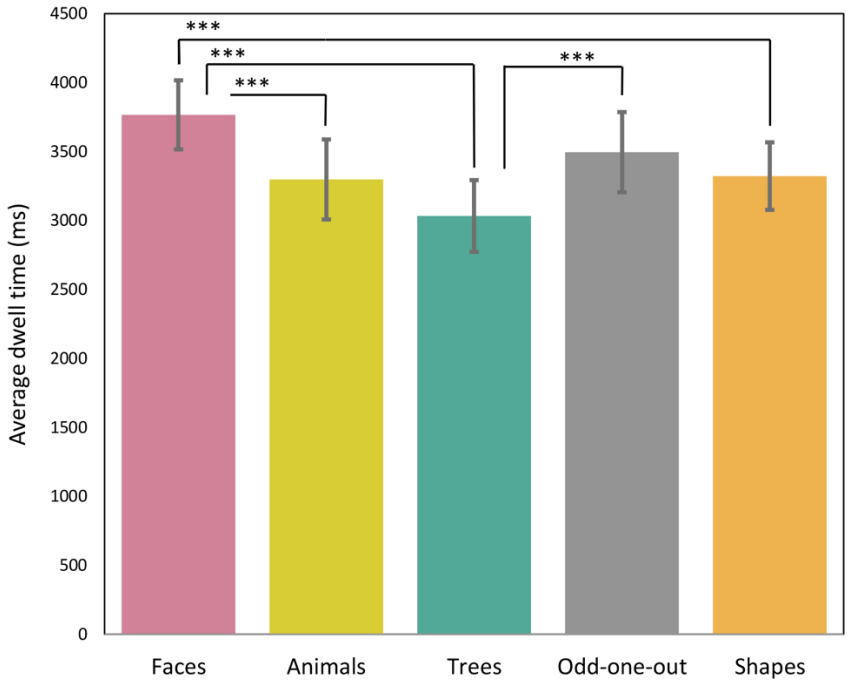


Figure 4.4. Bar chart showing average dwell time to five different pattern types across both age groups including the 95% CI error bars. Asterisk (*) indicating significant differences between pattern types (*** p<.001).

The variation of dwell time across pattern types also varied for infants and toddlers, with a significant interaction between pattern type and age, F (3.0, 119.84) =11.91, p<.001, η^2=0.06, BF_{10}=960382.88. Post-hoc Bonferroni corrected paired samples t-tests revealed that toddlers significantly spent more time than infants looking at patterns with animals (p<.001, BF_{10}=19404.89), trees (p<.001, BF_{10}=1552.09) and abstract shapes (p=.003, BF_{10}=14.12), but there were no dwell time differences for patterns with faces (p=.443, BF_{01}=3.27) and odd-one-out features (p=.385, BF_{01}=3.03), see Figure 4.5.

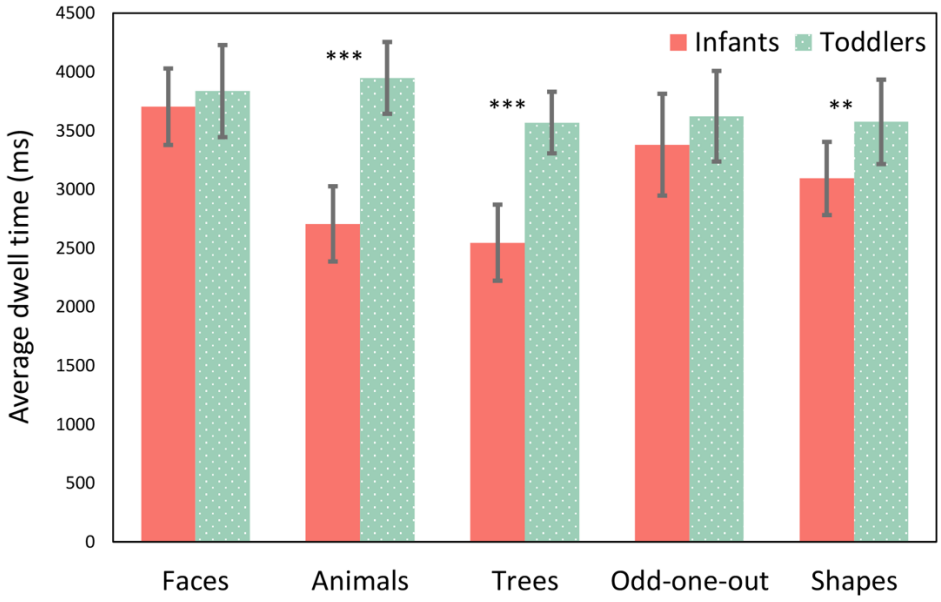


Figure 4.5. Bar chart showing average dwell time (milliseconds) and 95% CI error bars for each pattern type in infants and toddlers. Asterisk (*) indicating significant difference between infant and toddler dwell time to pattern types (**p <.01, *** p<.001).

These interactions were again further explored in each age group, firstly with post-hoc one-way repeated ANOVAs to confirm whether there was a significant difference in looking across pattern types and secondly with post-hoc Bonferroni

pairwise comparison tests to compare pattern types within each age group. A post-hoc one-way repeated measures ANOVA on infants revealed that pattern type had a significant effect on infant dwell time, F (2.43, 50.97) = 20.64, p<.001, η^2=0.50, BF_{10}=1068000000. Post-hoc Bonferroni corrected pairwise comparisons revealed that infants spent significantly more time looking at patterns with faces compared to patterns with animals (p<.001, BF_{10}=104001.51) and trees (p<.001, BF_{10}=11874.91). Infants also significantly spent more time looking at patterns with odd-one-out features compared to animals (p<.001, BF_{10}=101.41). In contrast, whilst pattern types had a significant effect on toddler's dwell time, F (4, 76) =3.44, p=.01, η^2=0.15, BF_{10}=3.41, further post-hoc Bonferroni comparison tests revealed that toddlers did not significantly differ in their looking between any pair of the five pattern types (p>.05). Although Bayesian paired samples t-tests reveal there is strong evidence to suggest that toddlers spend more time looking at patterns with animals compared to patterns with abstract shapes (BF_{10}=47.17) and trees (BF_{10}=15.85). All other Bayesian paired samples t-test comparisons revealed substantial to anecdotal evidence in support of the null hypothesis (BF_{01}<4.30), suggesting toddlers likely do not differ in their looking between other pairs of pattern types.

Do image statistics predict visual pattern preference in infants and toddlers?

To further investigate how the low-level and high-level image statistics of these patterns may affect looking in infants and toddlers, spatial and chromatic statistics were extracted from all patterns using MATLAB's Image Processing Toolbox (see Methods). The average dwell time was calculated for each of the 13 patterns for infants and toddlers. Four multiple regression models investigated whether the image statistics of the patterns predicted dwell time in infants and toddlers.

Model 1: Spatial image statistics

Figure 4.6 shows the correlation between spatial image statistics and average dwell time between infants and toddlers. Two multiple regressions were used to test whether the spatial statistics of the patterns (spectral slope, entropy, edge density and symmetry) could predict the average dwell time in infants and toddlers. Results indicated that spatial image statistics did not significantly predict a proportion of the variance of infant, R^2=0.42, *F* (4,8) =1.47, *p*=.30, BF_{01}= 1.87 or toddler average dwell time, R^2=0.24, *F* (4,8) =0.63, *p*=.65, BF_{01}= 3.82.

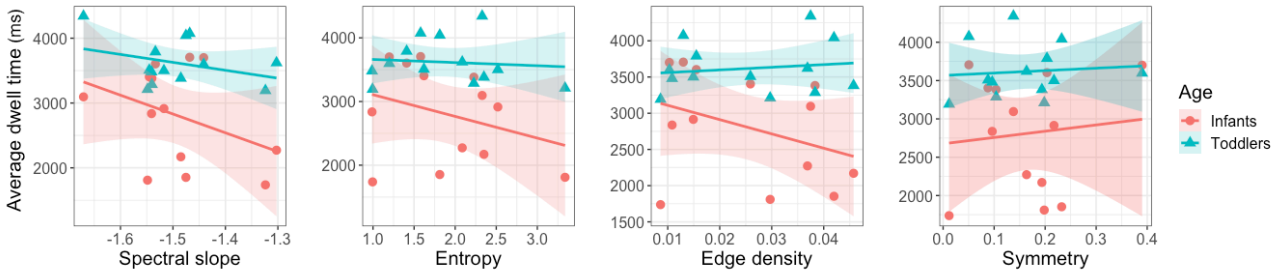


Figure 4.6. Scatterplots demonstrating the correlation between spatial statistics (spectral slope, entropy, edge density and symmetry) and average dwell time (ms) for infants and toddlers, with the 95% CI.

Model 2: Feature congestion

Figure 4.7 shows the correlation between the patterns' feature congestion and average dwell time between infants and toddlers. Although correlations were not significant, it appears infants may potentially look longer at patterns with lower feature congestion compared to toddlers who appear to look longer at patterns with greater feature congestion. Feature congestion was analysed in a separate multiple regression model as it's a high-level image statistic and an established measure of visual clutter (Rosenholtz et al., 2007). Two multiple regressions were used to analyse whether the pattern's feature congestion could predict looking time in infants and toddlers. Results revealed that feature congestion did not significantly predict a proportion of the variance for infant, R^2= 0.13, F (1,11) =1.68, p=.22. BF_{01}=1.30, and toddler average dwell time, R^2= 0.06, F (1,11) =0.66, p=.43. BF_{01}=1.77.

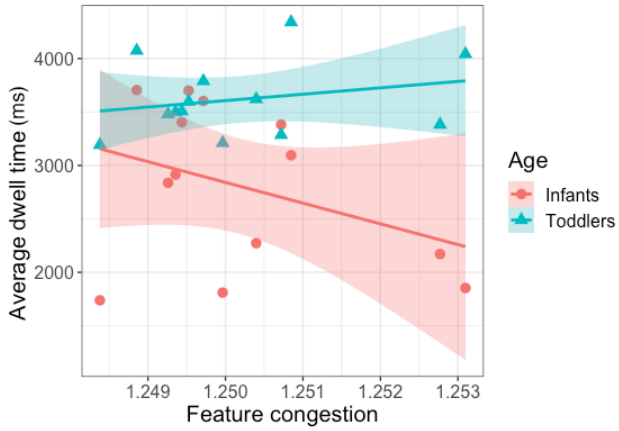


FIGURE 4.7. Scatterplot demonstrating the correlation between each pattern's feature congestion with average dwell time (ms) for infants and toddlers, with the 95% CI.

Model 3: Colour variance

Figure 4.8 shows the correlation between colour variance image statistics and average dwell time between infants and toddlers. Two multiple linear regressions were used to test whether the colour variance in patterns (SD hue, SD luminance, SD saturation) could predict the average dwell time in infants and toddlers. Results revealed that the SD hue, SD luminance and SD saturation were not significant predictors for infant, R^2=0.19, *F* (3,9) =0.72, *p*=.57, BF_{01}= 3.00, and toddler average dwell time, R^2=0.06, *F* (3,9) =0.19, *p*=.90, BF_{01}= 4.68.

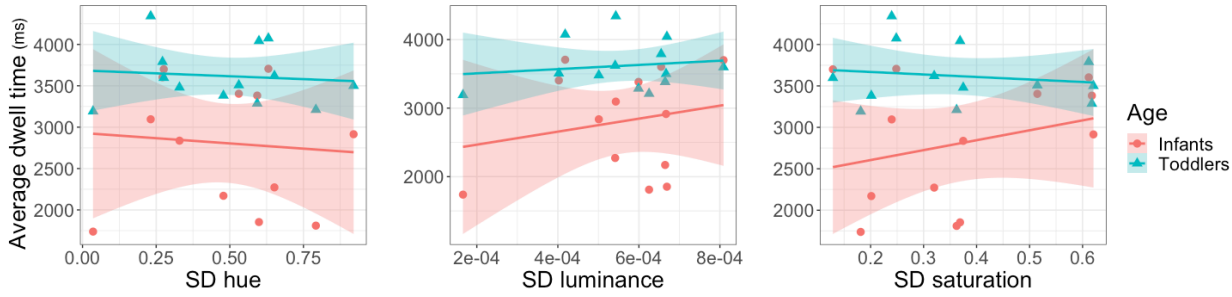


Figure 4.8. Scatterplots demonstrating the correlation between colour variance (SD hue, SD luminance and SD saturation) with average dwell time (ms) for infant and toddlers, with the 95% CI.

Model 4: Colour properties

Figure 4.9 shows the correlation between the colour properties and average dwell time between infants and toddlers. Two multiple regressions were used to test whether colour properties (mean saturation, mean luminance, and log axis ratio) could predict average dwell time in infants and toddlers. Results revealed that these colour properties did not significantly predict a proportion of the variance for infant, R^2= 0.20, *F* (3,9) =0.76, *p*=.55, BF_{01}=2.90, and toddler average dwell time, R^2=0.38, *F* (3,9) =1.82, *p*=.21, BF_{01}=1.41.

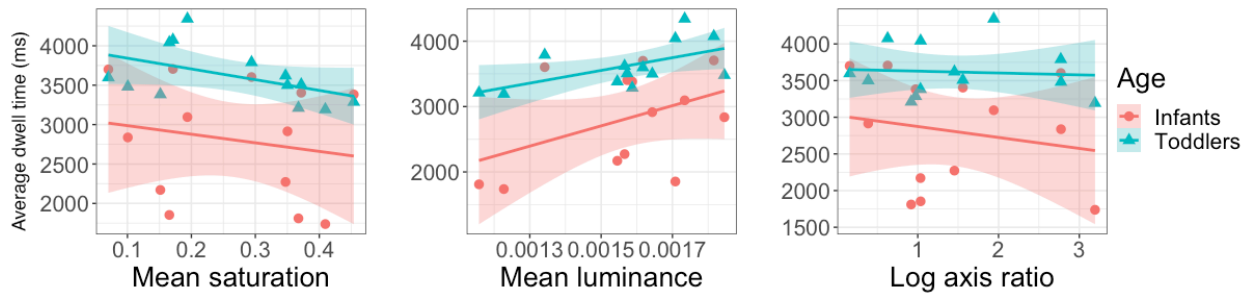


Figure 4.9. Scatterplots demonstrating the correlation between colour properties (mean saturation, mean luminance, log axis ratio) with average dwell time (ms), with the 95% CI.

Dominant chromatic hues

An alternative analysis approach was adopted for the image statistics of chromatic hues as the continuous variable of hue was non-normal. The most dominant hue was identified in each pattern (the hue with the most pixels) and each pattern was categorised into one of four dominant hue types (pink-reddish, charteuse-ish, cyan-blue-ish and achromatic hues) (see SI, Table S4). This converted the hue type from a continuous variable into a categorical variable of the pattern's most dominant hue type. The average dwell time to each dominant hue type was then calculated, enabling us to analyse whether the most dominant hue type in a pattern affected looking time in infants and toddlers.

A 4 x 2 mixed ANOVA with dominant hue type (within-subject levels: pink-reddish, charteuse-ish, cyan-blue-ish and achromatic hues) and age (between-subject levels: infants and toddlers) were analysed on average dwell time. Results indicated that regardless of age, the pattern's dominant hue type had a significant effect on average dwell time, F (2.31, 92.43) = 14.99, p<.001, η^2=0.12, BF_{10}=599201.64. Further Bonferonni corrected post-hoc tests indicated that looking was significantly longer for patterns with the most chartreuse-ish hues (p<.001, BF_{10}=3623.32) and the most achromatic hues (p<.001, BF_{10}=32.85) compared to patterns with the most pink-reddish hues. Significantly more time was also spent looking at patterns with the most chartreuse-ish hues (p<.001, BF_{10}=5112.98) and the most achromatic hues (p=.007, BF_{10}=44.12) compared to patterns with the most cyan-blue-ish hues.

See Figure 4.10. Age and a pattern's dominant hue did not significantly interact, F (2.31, 92.43) =2.85, p=.055, η^2=0.02, BF_{10}=1.41.

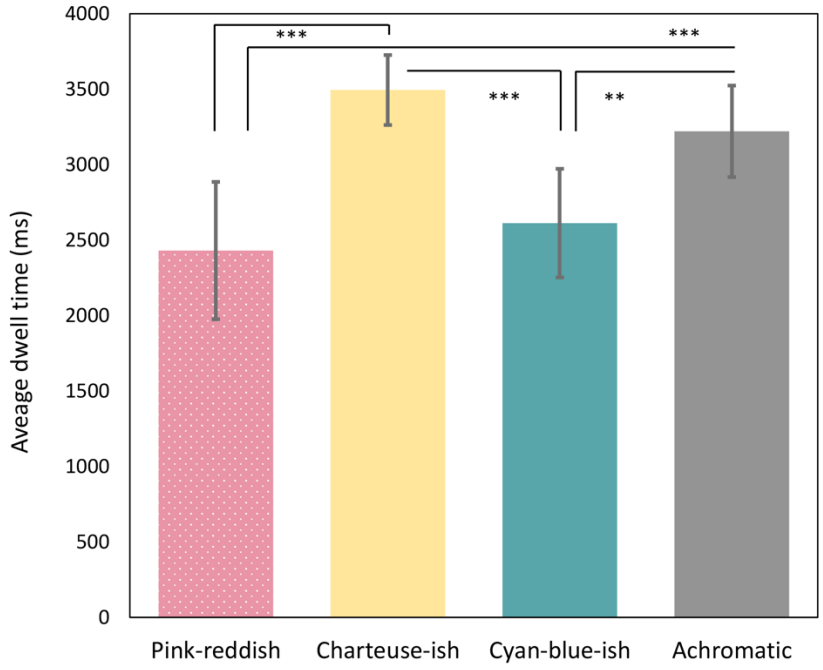


Figure 4.10. Bar chart showing significant differences in average dwell time (ms) for patterns with different dominant hues (pink-reddish, chartreuse-ish, cyan-blue-ish, achromatic), with the 95% CI error bars. Asterisk (*) indicating significant difference between (**p <.01, *** p<.001).

4.5 Discussion

The current study explored visual pattern preferences in 6-8 month old infants and 2-3 year old toddlers. Toddlers looked longer at the patterns than infants, and particularly looked longer than infants at patterns with trees, abstract shapes and animals. Infants appeared to have a preference for patterns with schematic faces and odd-one-out features and looked at an equivalent amount of time at those patterns compared to toddlers. The dominant hue of the pattern

also affected how long infants and toddlers look, with longest looking for patterns with chartreuse or achromatic dominant hues. However, a range of chromatic and spatial image statistics did not predict looking time in either infants or toddlers. We discuss each of these findings in order.

First, the finding that visual pattern preferences appear to differ for 6-8 month old infants and 2-3 year old toddlers gives insight into visual and perceptual development. Although toddlers spend more time looking at the patterns than infants, when the patterns are categorised into pattern types, infants spend a similar amount of time looking at patterns with schematic faces and odd-one-out features as toddlers, suggesting the relative importance of these features in infant pattern perception. The early preference for patterns with faces and odd-one-out features is likely due to the way in which these features automatically grab attention, even in infancy (Adler & Orprecio, 2006; Frank et al., 2009, 2014; Gliga et al., 2009; Kwon et al., 2016; Quinn & Bhatt, 1998). Infants' preferences for patterns with faces and odd-one-out features can be linked to theories of pre-attentive automatic visual processing (Treisman et al., 1992; Treisman & Gelade, 1980), and established infant research showing biased preference for faces and colour pop-out (Catherwood et al., 1996; Davies & Franklin, 2002; Frank et al., 2009, 2014; Gliga et al., 2009; Kwon et al., 2016). However, it is also important to note that Bayesian analyses revealed that there is anecdotal evidence to support the null hypothesis, suggesting that we cannot be confident that there are no looking time differences between infants and

toddlers when viewing patterns with faces and odd-one-out features, thus we may need to interpret these findings with caution. Furthermore, interestingly, post-hoc analyses indicated toddlers did not show a stronger visual preference to patterns with faces and odd-one-out features than other features such as abstract shapes, trees and animals. This is a surprising finding considering that past research has indicated children and adults show biases to faces and odd-one-out features (Adler & Orprecio, 2006; Gerhardstein & Rovee-Collier, 2002; Kadooka & Franchak, 2020), although it suggests that other visual features of patterns can be just as important to 2-3 year old toddlers.

Second, the age-related looking differences between patterns with animals, trees and abstract shapes may reflect developmental attentional changes in bottom-up and top-down processing with age (Franchak et al., 2016; Frank et al., 2014; Tummeltshammer & Amso, 2018). Toddlers may be using more top-down endogenous attention processing to select and inhibit where to attend in the patterns compared to infants who may be more likely to visually process the patterns using bottom-up exogenous attention processing. For example, patterns with animals like ‘Hop To It’ and ‘Fox Tale’ have small animals embedded in the pattern. Toddlers may be using their semantic knowledge about animals to visually search for the rabbit or fox in the pattern compared to infants who may not have yet gained the semantic knowledge to understanding what animals are. These inferences also support previous developmental studies which have shown that toddlers and children use scene-context and

semantic content to guide their looking (Helo et al., 2017; Rider et al., 2018), and as children get older, they get better at learning to prioritise where they want to look at (Franchak, 2020; Kadooka & Franchak, 2020).

Third, the influence of the dominant hue of a pattern on both infants and toddlers is a novel result. Previous work has investigated infant and child preferences for hues when presented as a single colour patch (Brown & Lindsey, 2013; Franklin et al., 2008; Skelton & Franklin, 2020). However, research has not previously investigated how hue influences visual preference for varied multicoloured patterns. Here, both infants and toddlers prefer patterns with the most chartreuse-ish hues and the most achromatic hues compared to patterns with most pink-reddish and most cyan-blue-ish hues. This is the reverse pattern to infant visual preference for single patches of colour, where infants look longest at blue and least at chartreuse (Bornstein, 1975; Franklin et al., 2008, 2010; Skelton & Franklin, 2020; Zemach & Teller, 2007). It seems unlikely that this result was influenced by other image statistics, such as the spatial image statistics of the patterns in the dominant hue groups, as the average looking time of the dominant hue was calculated and no significant findings were found from previous individual image statistics analyses. This finding suggests that hue preferences vary according to whether the hue is distributed across a multielement ‘scene’ (e.g., a multicoloured abstract pattern) or an object (e.g., a single homogenous stimulus). Many of the patterns with chartreuse-ish dominant hues had a yellowish background, and one possible

tentative explanation is that there is a visual bias for yellowish backgrounds that mimic the hues of natural daylight (people may also most commonly choose a yellowish white paint for their walls for this reason (A. Franklin & Dulux Ltd., personal communication, December 2016). These findings may help inform designers when selecting a dominant colour scheme for their designs optimized for infants and young children.

Finally, our analyses do not find that the low-level and high-level image statistics of the abstract patterns predict visual preference in infants and toddlers. However, our Bayesian multiple regression analyses indicate there is generally anecdotal or a lack of evidence in support of the null hypothesis for all image statistics, indicating that firm conclusions about a null effect of image statistics on looking cannot be made. The insensitive Bayes factor here is likely be due to only having 13 patterns, and perhaps also the type of patterns used which had a restricted range for some of the image statistics variables. A larger set of more variable patterns which more appropriately capture the full range of each image statistic might reveal significant predictors. However, the current study makes a useful contribution by establishing an analysis template for investigating whether chromatic and spatial image statistics predict looking in early development. One limitation with the existing image statistics toolbox that we developed is that it is not currently specifically designed to reflect infant vision. Future studies could improve the toolbox to address this and may want to consider additional image statistics not considered in our analyses which may

affect visual preference, such as fractal lacunarity (Aks & Sprott, 1996; Mori et al., 1996). It would be interesting for further developmental research to see whether these image statistics predict looking preferences for natural scenes to further explore potential tuning to natural image statistics during visual development.

Overall, this study has progressed understanding of the developmental changes in visual pattern preference**.** Our findings further contribute to understanding the development of visual perception and attention. The findings also have implications for those in industry aiming to design patterns or products that are optimized for infant and child vision. More broadly, further research in this area has potential to contribute to theoretical debate on the extent to which visual development tunes into statistical regularities, and the role of image statistics in aesthetic preference.

Chapter 5

Paper 4: Infants do not look longer at abstract images with natural chromatic scene statistics

Teresa Tang[1], Alice Skelton[1], Jenny Bosten[2], James Alvarez[3] & Anna Franklin[1]

[1] The Sussex Colour Group, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

[2] Sussex Vision Lab, School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

[3] School of Psychology, University of Sussex, Falmer, BN1 9QH, UK

5.1 Abstract

The adult visual system is optimally tuned to the statistical regularities of natural scenes which are known as natural scene statistics (Simoncelli, 2003; Simoncelli & Olshausen, 2001). Natural scenes have chromatic statistics, for example, they have a distinctive chromatic distribution that varies predominantly along the blue-yellow daylight axis in colour space. This scene statistic is reflected in both colour perception and aesthetics: both adults and infants are less sensitive at discriminating colours along the blue-yellow daylight axis (Bosten et al., 2015; Skelton et al., 2021), and adults also rate art or abstract images with chromatic distributions that align with the blue-yellow daylight axis as more aesthetic (Juricevic et al., 2010). Here, we investigate whether infant preferential looking is sensitive to the naturalness of the chromatic distributions of abstract images. Using eye-tracking, 6–8 month old infants (N=36) were shown chromatic Mondrian images varying along eight colour axes. Results

revealed that infants look equally at the images regardless of the natural chromatic distribution. These findings could suggest that even though infant hue sensitivity is aligned with the natural chromatic distribution of scenes at 4-6 months old, this alignment is not seen for infants' visual preferences at 6-8 months. The implications for understanding how the perceptual system tunes into the visual world during development are discussed.

Keywords: natural scene statistics, chromatic image statistics, visual development, visual preference, colour, developmental aesthetics

5.2 Introduction

Natural scenes are real-world scenes that are typically images of the natural environment such as mountains and trees (Olshausen & Field, 2000), but they can also include human-made environments like buildings and other human-made objects (Geisler, 2008). Natural scene statistics refer to the statistical properties and regularities of natural scene images, such as chromatic or spatial regularities. It has been argued that our visual system has evolved to respond to the statistical regularities in our natural world (Geisler, 2008; Olshausen & Field, 2000; Simoncelli, 2003). For example, when natural images are decomposed into their spatial components, there is a characteristic relationship between the luminance contrast and spatial frequency where the contrast increases as spatial frequency decreases (G. Cole & Wilkins, 2013). When this

is plotted, natural images show a distinct $1/f^{\alpha}$ (or equivalently $f^{-\alpha}$) spatial frequency, where α (the amplitude spectral slope) is approximately 1 (Burton & Moorhead, 1987; D. Field & Brady, 1997; Hansen & Hess, 2006; Olshausen & Field, 2000; Simoncelli & Olshausen, 2001; Tolhurst et al., 1992; Torralba & Oliva, 2003). These findings suggest the visual system may be optimally sensitive for spatial 1/*f* statistics because they reflect the spatial properties of our natural environment (D. Field, 1994; O'Hare & Hibbard, 2011).

It has also been suggested that there is a relationship between natural scene statistics and aesthetics, where the characteristic statistical properties of natural scenes can contribute to their perceived aesthetic value (Kardan et al., 2015; Redies, 2007; Redies et al., 2007; Spehar et al., 2003). Many studies have found a preference for images with a natural 1/*f* spatial frequency (Isherwood et al., 2021; Nguyen & Spehar, 2021; Spehar et al., 2003, 2016; Spehar & Taylor, 2013), and when images deviate from this natural 1/*f* spatial frequency, it creates visual discomfort and decreases perceived aesthetic value (Fernandez & Wilkins, 2008; Hibbard & O'Hare, 2015; Juricevic et al., 2010; O'Hare & Hibbard, 2011; Penacchio & Wilkins, 2015). This may reflect the underlying neural mechanisms of aesthetic perception, where efficient coding of sensory information by the brain may be related to aesthetic perception (Redies, 2007; Spehar et al., 2015). Hence, the argument is that images with a different spatial structure to natural images may cause visual discomfort as they are not optimised to the biological mechanisms of the visual system.

Similar to spatial properties, the chromatic statistics of natural scenes also have unique and characteristic properties (Atick et al., 1993). Previous research has found that the average colour distribution of natural scenes tends to vary along the blue-yellow daylight axis, an axis in colour space that extends from colours that appear bluish to yellowish (Burton & Moorhead, 1987; McDermott & Webster, 2012; Shevell & Kingdom, 2008; Van Hateren & Ruderman, 1998; Webster et al., 2007; Webster & Mollon, 1997). For example, if we were to plot all the chromaticities that appeared in a natural image, we would find that on average the distribution of these colours will fall predominately along the axis in colour space which extends from blue to yellow. This blue-yellow daylight axis closely corresponds to the daylight locus, a line through colour space which depicts the chromaticities of the different variations of natural daylight throughout the day (Granzier & Valsecchi, 2014; Panorgias et al., 2012; Shepard, 1992). The daylight locus is also associated with poorer colour discrimination along the blue-yellow daylight axis in adults (Álvaro et al., 2017; Bosten et al., 2015; Krauskopf & Gegenfurtner, 1992; Pearce et al., 2014). This poor discrimination potentially occurs because the blue-yellow daylight axis has the greatest variation of surface reflectance and illumination as it closely corresponds to the daylight locus (Bosten et al., 2015; Delahunt & Brainard, 2004). Hence, having poorer colour discrimination performance along the blue-yellow daylight axis may be related to the statistical chromatic regularities of the environment and could be an indication of being optimally tuned to the chromatic statistics of natural scenes. Recently, infants as young as 4 months

old have also been found to have poorer blue-yellow colour discrimination (Skelton et al., 2021). This finding suggests that colour vision is either optimised to the chromatic statistics of natural scenes at an early age, learnt very rapidly in the first few months of life, or is innate.

It has also been suggested that the naturalness of an image's chromatic distribution affects visual discomfort and aesthetics (Juricevic et al., 2010). For example, artists implicitly produce paintings which closely follow the blue-yellow chromatic statistics of natural images (Montagner et al., 2016) and observers prefer original paintings compared to colour-manipulated versions (Albers et al., 2020; Altmann et al., 2021; Nakauchi et al., 2018; Nascimento et al., 2021). In another study, Mondrian images (Land, 1983) showing rectangles varying randomly in size and location, were presented to adults with the chromaticities of the rectangles distributed along different colour axes (Juricevic et al., 2010). For example, one image consisted of varying colours distributed along the natural blue-yellow daylight axis and another image consisted of varying colours distributed along the orthogonal "unnatural" red-green colour axis. It was found that images with an average chromatic variation around the natural blue-yellow daylight axis were rated lowest in visual discomfort and highest for artistic merit by adults. These findings suggest that the visual system is optimally tuned to the blue-yellow characteristics of our environment and that this is reflected in aesthetics and visual discomfort.

Here we ask whether given the findings outlined above, natural chromatic scene statistics are also reflected in infants' early visual preferences. Little is known about whether the visual system tunes into natural image statistics during development, or whether this has occurred during evolution. The demonstration that infant hue sensitivity relates to the natural blue-yellow distribution of colours in natural scenes (Skelton et al., 2021) is the earliest example of infant vision reflecting natural chromatic scene statistics. Another study has found that infants appear to be neurally sensitive to the texture statistics of natural scenes by 9 months of age (Balas et al., 2018), but it has also been found that children are not optimally sensitive to the spatial 1/*f* spatial statistics of natural images until 10 years old (Ellemberg et al., 2012). For colour, given that both adults and infants are less sensitive at discriminating colours along the blue-yellow daylight axis and adults rate chromatic distributions that align with this natural axis as most aesthetically pleasing (Álvaro et al., 2017; Bosten et al., 2015; Juricevic et al., 2010; Krauskopf & Gegenfurtner, 1992; Pearce et al., 2014; Skelton et al., 2021), it raises the question whether chromatic scene statistics are reflected in infant visual preferences as well.

By as young as 4 months, infants are known to have visual preferences for looking longer at some colours, such as blue, than others such as chartreuse (Franklin et al., 2008, 2010; Skelton & Franklin, 2020). How long an infant looks at a stimulus is not a direct measure of how much they like it (C. Taylor et al., 2013), however, studies have highlighted similarities between infant looking and

adult liking. In particular, infants' visual preferences for single colours significantly relates to how much adults like those colours (Skelton & Franklin, 2020), potentially suggesting that adults' aesthetic response is rooted in infants' early visual preference. Whilst we know about infants' visual preference for single colours, it is unknown how infants respond to multicoloured images with chromatic distributions, and whether infants prefer to look at those with natural blue-yellow chromatic distributions that reflect natural scenes. The current study investigates this using similar Mondrian stimuli to the study which demonstrated that adult visual aesthetics and discomfort for chromatic Mondrian images relates to natural chromatic statistics (Juricevic et al., 2010). The use of chromatic Mondrian images isolates the chromatic statistics from other statistical and semantic features of natural images. We show 6–8 month old infants these Mondrian images and record their eye-movements to establish whether infants have a visual preference for natural chromatic images that vary closely along the blue-yellow daylight axis compared to "unnatural" chromatic images which vary along different colour axis directions. We predict that infants will look longer at natural chromatic Mondrian images varying along the natural blue-yellow daylight axis compared to "unnatural" chromatic Mondrian images which vary along different colour axis directions. The study will enable us to further understand the extent to which infant vision is aligned with the statistics of natural scenes, and the relationship between infant visual preference and adult aesthetics.

5.3 Methods

Participants

Forty-one 6–8 month olds (25 females, *M*= 7.18 months, *SD*= 0.53) were tested. One infant was excluded due to insufficient trials and 4 additional infants were excluded as they had a family history of colour vision deficiency. Data from a total of 36 infants were analysed (22 females, *M*= 7.24 months, *SD*= 0.50). Infants were full term and had a minimum birth weight of 2500g.

Stimuli

As in Juricevic et al. (2010), eight types of images were generated by an algorithm which produced overlapping rectangles of random size, location and varying colour distributions along a single colour axis. Each of the eight images followed eight different colour axes at 0- 180°, 22.5- 202.5°, 45- 225°, 67.5- 247.5°, 90- 270°, 112.5- 292.5°, 135- 315° and 157.5- 337.5° in the MacLeod-Boynton chromaticity diagram (see Figures 5.1-5.2 for examples of Mondrian images in each colour axis). The MacLeod-Boynton chromaticity diagram is a physiological colour space which represents the visual system's response to LMS cone excitations (MacLeod & Boynton, 1979) based on Smith & Pokorny (1975) cone fundamentals. The colour contrasts were based on a variant of the MacLeod-Boynton chromaticity diagram and the Smith & Pokorny (1975) cone fundamentals were scaled to approximately equate chromatic variations along the axes according to the following equations: L/(L+M) = 1955 (r_{mb} – 0.6568),

S/(L+M) = 5533 (b_{mb} – 0.01825), Luminance = 70Lum, where r_{mb}, b_{mb}, are the chromaticity coordinates in the MacLeod-Boynton colour space; 0.6568, 0.01825 are the *r*, *b* coordinates of Illuminant C (a standard illumination that represents daylight, Kránicz, 2016) and Lum corresponds to the luminance relative to the mean luminance (42 cd/m^2) in the image [$(I-I_{mean})/I_{mean}$]. The scaling of the Smith and Pokorny (1975) cone fundamentals is important so that the spectrum locus peaks at S/(L+M) =1 and the L+M matches the luminosity function. Images were isoluminant across all trials. The size of the rectangles within each image randomly differed in the range of 40 to 140 pixels (visual angle range of 0.45° to 1.58°). The saturation value of each rectangle was randomised between 0 and 0.0156 radians, the maximum saturation available in the gamut. The size of the stimuli and their components were selected to ensure infants could see the stimuli based on infants' visual acuity (Teller et al., 1986).

Images were placed on a grey background and displayed on a 22-inch Diamond Pro 2070SB CRT monitor (Mitsubishi, Tokyo Japan) (screen resolution 1600 x 1200 pixels at 85 Hz refresh rate) controlled by a visual stimulus generator, ViSaGe (Cambridge Research Systems, Rochester, UK). On the whole monitor display, the stimulus was presented with a visual angle of 10° (screen dimensions= 495 x 484.5 mm, eye to screen distance= 600mm). Eye movements were recorded with an EyeLink 1000 Plus eye-tracker (SR

Research, Ontario, Canada) and the experiment was run in MATLAB (R2016a, MathWorks Inc.).

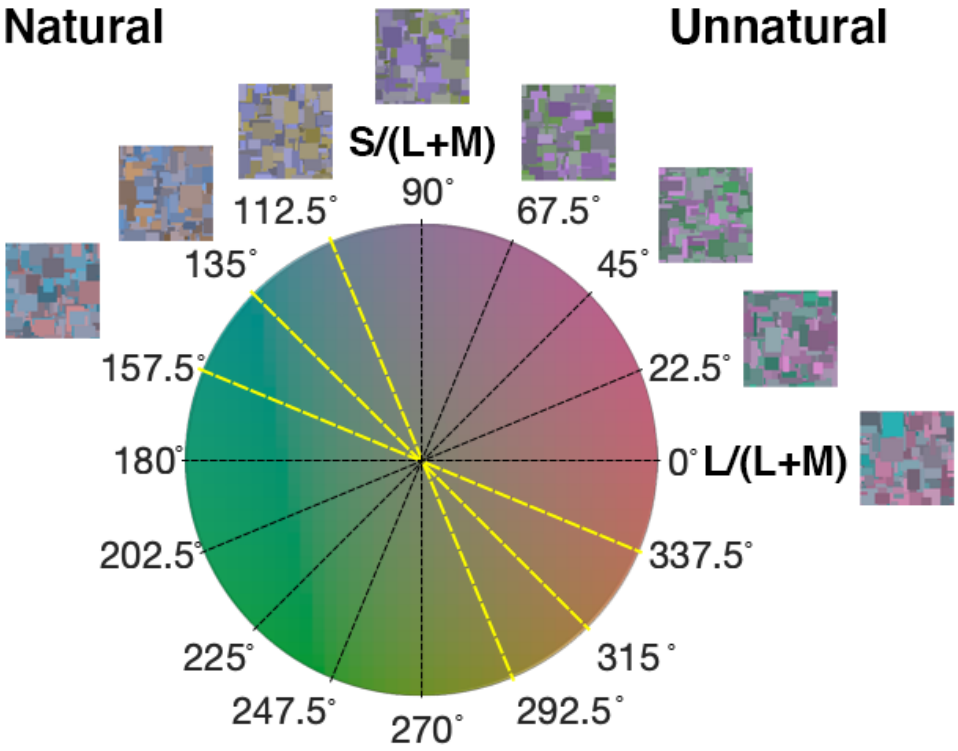


Figure 5.1. Examples of all eight types of abstract Mondrian images represented in angles (°) in polar space in the MacLeod-Boynton chromaticity diagram. The MacLeod-Boynton chromaticity diagram is an isoluminant physiological colour space which represents the visual system's response to L, M and S cone excitations (MacLeod & Boynton, 1979). The figure shows the cardinal directional axes of colour vision, with 'L/(L+M)' and 'S/(L+M)'. Each image type is created by producing random varying colours of rectangles of random size and locations along a single colour axis in MacLeod-Boynton colour space. For example, the image at 0° consists of colours randomly varying along the 0-180° colour axis. Images along the marked dashed yellow line in colour space (112.5- 292.5°, 135-315° and 157.5- 337.5°) are varying colour distributions closest to the natural blue-yellow daylight axis, and thus would be considered as images with natural colour variation. In contrast, all other images on the opposite colour axes (i.e., at 0-180°, 22.5-202.5°, 45-225°, 67.5-247.5°, 90-270°) would be considered as images with "unnatural"

colour variation. Please note that as the size of rectangles were randomised for every trial, a different Mondrian image was produced each time, but each image would have the same chromatic distribution along each colour axis.

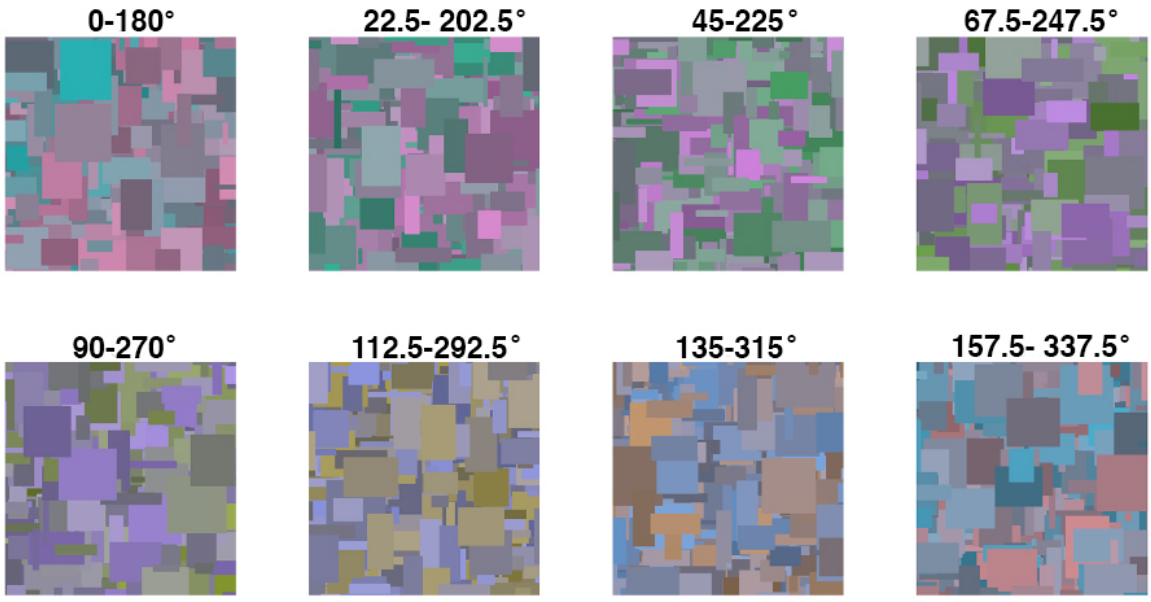


Figure 5.2. Examples of the eight types of chromatic Mondrian images. Angles written above each image show the defined colour axis of each image. For example, image 0-180° represents different colour pattern distributions varying along the 0-180° colour axis. Please note again that as the size of rectangles were randomised for every trial, infants would see a different image each time, but each image would have the same chromatic distribution along each colour axis.

Design and procedure

Infants were seated in a car seat secured on a chair approximately 60cm away from the screen. Infant looking was calibrated using a 4-point calibration procedure, where infants fixated on a small looming black and white motion target on various locations of the screen. Using a two-alternative forced choice paired comparison procedure via eye-tracking, each infant was presented with

two different images, presented simultaneously side by side on the screen in each trial (see Figure 5.3). This led to a maximum of 64 trials for all eight images to have been presented on the screen at least once and counterbalanced on both sides of the screen. Each trial was presented for 4 seconds to ensure infants had time to process the images. When the participant fixated on one of the two images, the next trial would be presented. After each trial a visual attention grabber was presented on the screen to maintain the participant's gaze at the centre of the screen. The attention grabber was presented until the infant's gaze fixated on the attention grabber. The visual attention grabber was a looming black and white motion target played with a melody (see Figure 5.3). The trials continued until the infant became disengaged.

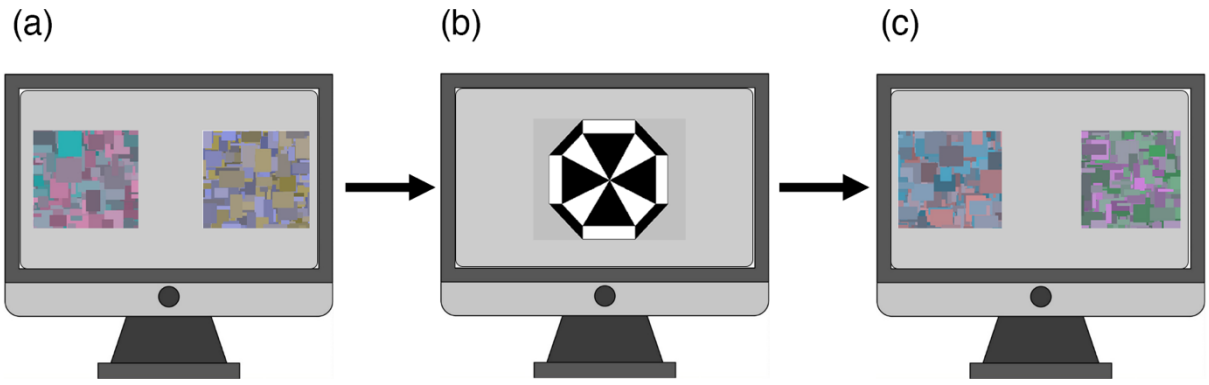


Figure 5.3. Example of the procedure setup with **(a)** one trial showing two types of Mondrian images presented side by side, **(b)** interstimulus interval with the visual attention grabber and **(c)** the following trial of two types of Mondrian images.

5.4 Results

On average, each infant completed 111.61 trials (SD=54.87, minimum= 30 trials, maximum= 244 trials). A trial was considered valid if infants looked at the screen for a minimum of 0.50 seconds. There was no significant effect on the number of trials completed and the colour axis, F (7, 245) =0.68, p=.691, η^2=0.02.

Eye-movement data were converted to be processed in MATLAB (R2016a, MathWorks Inc.) to extract looking time. The looking time was extracted from a specified region of interest in the images with an additional 10% of pixels of the image pixel size for margin of error.

The average looking time to each image was calculated (total time spent looking at the image across all trials/total number of trials completed for each infant for that image). We can see in Figure 5.4a and Figure 5.4b that the average looking time per trial to each of the eight images does not vary for all infants in all eight colour axes in cartesian and polar space. However, we can see that individually, some infants appear to have a visual preference for images with natural chromatic distribution and other infants appear to have visual preferences for images with “unnatural” chromatic distribution (see Figure 5c). A repeated measures ANOVA revealed there was no significant main effect of colour axis on looking time for infants, F (7, 245) =0.89, p=.515, η^2=0.03. A

Bayes factors was also additionally computed to predict the ratio of the probability of the data under the null hypothesis compared to the alternative hypothesis, as null hypothesis significance tests provide no evidence about the strength of the null hypothesis (Dienes, 2014). Using a multivariate Cauchy prior distribution with a fixed effects scale parameter of 0.5 (Rouder et al., 2012), a post-hoc Bayesian repeated measures ANOVA indicated there was very strong evidence in favour of the null hypothesis (BF_{01}=44.82), suggesting that the data is approximately 44.82 times more likely to occur under the null hypothesis (no effect of colour images on average looking time) than the alternative hypothesis.

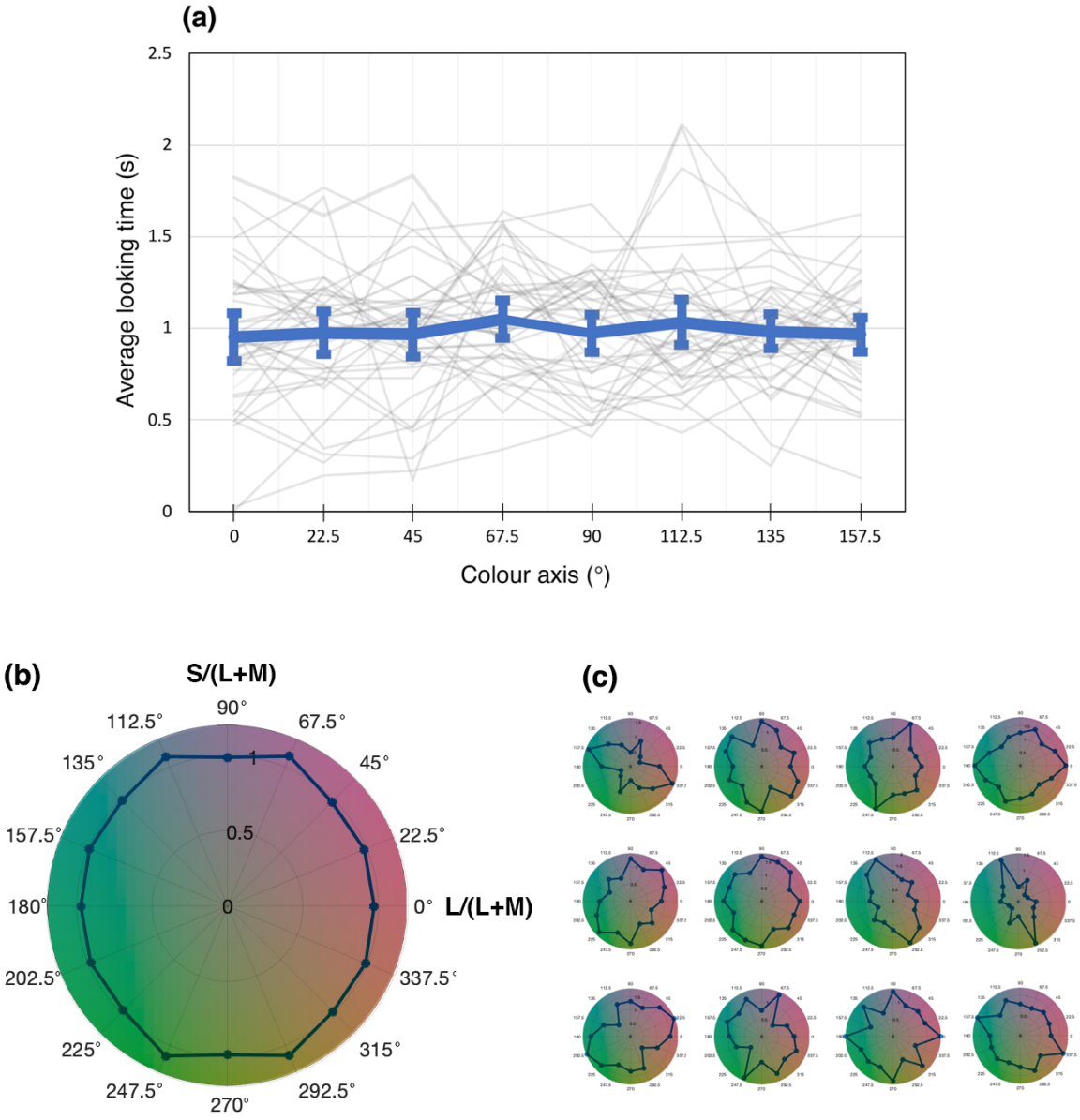


Figure 5.4. (a) The average looking time (in seconds) to each colour axis image plotted in cartesian space on a line graph, with each infant's looking time marked in grey and the average looking time marked in bold blue with the 95% CI error bars; **(b)** shows the same average looking time to each colour axis image plotted in polar space of the MacLeod-Boynton chromaticity diagram; and **(c)** shows individual average visual preferences to each colour axis plotted in MacLeod-Boynton colour space. In polar space (Figures 5.4b-5.4c), looking time is plotted as the distance from the origin, hence the further away from the centre origin, the longer the looking time. Please note that points 180° apart are from the same colour axis and will have

the same looking time. For example, the looking time at 0° is the same as the looking time at 180°. Figure 5.4c shows examples of individual infants looking to each colour axis. Each ellipse represents the infant's preferred average chromatic distribution, i.e., their average looking preference to every Mondrian image type. For example, if the ellipse is elongated towards a certain colour axis, it indicates greater looking for that specific colour axis, for e.g., the ellipses of two infants in the third and fourth column of the second row, show these infants had a stronger visual preference for natural chromatic images, as the looking time ellipse elongates at 112.5-292.5°.

5.5 Discussion

The current study aimed to establish whether infants look longer at Mondrian images that have a natural blue-yellow chromatic distribution. Individually, infants show varied responses, with some infants having distinct visual preferences for images with natural chromatic distributions, and other infants showing visual preferences for "unnatural" colour distributions (see Figure 5.4c). However, on average, 6–8 month old infants spent an equal amount of time looking at all images irrespective of their chromatic distributions, and did not show a visual preference for Mondrian images with natural chromatic scene statistics.

Our results contrast with the finding that infants' hue sensitivity aligns with the natural blue-yellow chromatic distribution at 4-6 months (Skelton et al., 2021). There are three possible explanations for this potential discrepancy in findings. First, it is possible that infants are not tuned to the chromatic scene statistics at

4-6 months old and the finding on hue sensitivity was not reliable and was a spurious result. We would be surprised by this given the psychophysical rigour of Skelton et al's (2021) study, yet despite this replication is always important. Second, it is possible that infant colour vision is tuned to natural chromatic scene statistics in early infancy, but it is only for hue sensitivity and visual preference is not affected. For example, one hypothesis is that visual preference takes longer to tune into natural scene statistics than low-level discrimination. Potentially, it could also be that natural chromatic scene statistics do not affect visual preference but start to have an influence when a more mature aesthetic judgement starts to develop (e.g., at around 2 years (Lobue & Deloache, 2011)). Little is known about the development of visual aesthetics, and there could be great potential in applying a developmental approach to the field of visual aesthetics to further understand why certain stimuli are visually appealing.

Third, it is possible that the current study was not sensitive enough to reveal an effect. Statistically we can be confident of the effect, as the Bayes factor indicates strong evidence in support for the null hypothesis. However, on average, infants do not look for very long at the stimuli (each stimulus is looked at on average for 1 second during a 4 second trial), which could suggest that the Mondrian stimuli were not sufficiently engaging for infants. Although we verified that the chromaticities of the Mondrian images were discriminable for infants at the age we tested, some colours were very desaturated and there is a

possibility that when each colour is shown as a small patch along with other colour patches that this affects discriminability. Previous research has indicated that how quickly infants orient to a stimulus is related to salience (attention-getting), whereas how long an infant looks at a stimulus is related the novelty and complexity of the stimulus (attention-holding) (Cohen, 1972; Kwon et al., 2016). Hence, it is possible that the lack of complexity in our chromatic stimuli accounted for this short looking duration. Another possibility is that infants' attention was focused on the spatial elements of the stimuli (e.g., the edges or arrangements of the rectangles that form the Mondrian) and that this distracted them from encoding the chromatic distribution of the stimuli. Regular multicoloured checkerboards or phase-scrambled stimuli with varying chromatic distributions may help make the chromaticity of the stimuli more salient.

Given the uncertainty on how to interpret the null effect and how to distinguish between the three potential explanations outlined above, at this stage firm theoretical implications for our finding cannot be specified. However, further research could investigate the questions raised by the current study which if answered would have implications for fundamental debate on several issues such as the role of experience in perceptual development (Bornstein et al., 2011), efficient encoding (D. Field, 1994; Simoncelli, 2003; Simoncelli & Olshausen, 2001) and the nature of visual aesthetics (Graham & Redies, 2010). Further research could investigate whether infants prefer natural scenes or art with natural chromatic distributions as opposed to scenes or art where the

distribution of chromaticities has been manipulated (Nascimento et al., 2021). Future research could also investigate the relationship between other chromatic and spatial natural image statistics and infant visual preference. Recently infant head-mounted cameras have been used as a tool to capture images from infants' visual environment (Jayaraman & Smith, 2020; L. Smith et al., 2011, 2015). This could be further explored by measuring how other image statistics of infants' visual world change with development and the visual implications of this on development.

In summary, this study is the only study to date to investigate whether infants' visual preferences reflect the chromatic statistics of natural scenes. We do not find an effect at 6-8 months old. Further research can distinguish between the specified explanations for this lack of effect. Doing so would further contribute to our understanding of whether the visual system develops to become optimised to the natural scenes of our environment, and whether this optimisation also contributes to the development of visual aesthetics.

Chapter 6

6.1 Overall contribution

The four papers of this thesis make a number of theoretical, methodological and applied contributions to science, industry and society. These are outlined below.

Theoretical contributions

One aim of this thesis was to make a theoretical contribution to the fields of perceptual development, colour perception and vision science. The thesis has achieved this in a number of ways. First, the thesis contributes to further understanding of visual development. We demonstrate age-related changes to pattern preferences in infants and toddlers, which potentially highlight developmental changes to perceptual attentional processing in infancy, such as more bottom-up attentional processing in infancy but more top-down attentional processing by 2-3 years old. Second, the thesis contributes to the theory that proposes that colour perception is aligned with natural scene statistics. We learn that 6–8 month old infants do not show a visual preference to natural chromatic scene statistics which instigates new questions about how and why colour vision is tuned to the chromatic scene statistics at 4 months old (Skelton et al., 2021) but not shown with infant visual preference of abstract Mondrian images. This finding also raises the question of how chromatic scene statistics contribute to the aesthetic preferences in adults and raises new questions about the role of visual preferences in infancy in the development of an aesthetic

response. Finally, the findings of the thesis have highlighted the importance of colour in infants' and children's daily life. For instance, the thesis shows the importance of the dominant colour of a pattern on infant and toddler looking. Also, the thesis demonstrates that older children and adolescents with CVD report more difficulty with colour-related tasks and activities used in education than children without CVD.

Methodological contributions

The thesis makes a number of methodological contributions to developmental vision science. The thesis presents the first tablet-based optometric test which uses gamification and psychophysics to diagnose CVD in young children from 4 years. The automatic model-specific colour calibration enables online remote testing which is particularly valuable currently during the current COVID-19 pandemic with risks associated with in-person optometric assessments. *ColourSpot*'s source code, interface, animation and other elements of the test is available to download for research purposes. As well as diagnosing CVD, the development of an accurate gamified psychophysical test helps measure discrimination threshold estimates in children and *ColourSpot*'s methods could be adapted for efficient, effective and remote testing of other visual abilities in children. This research can help a number of practitioners and scientists including optometrists and ophthalmologists seeking a quick and accurate paediatric diagnostic test of CVD, as well as those interested in colour calibration of tablet devices, and other vision scientists interested in applying

gamified psychophysical methods to young children. In addition, the thesis provides an image statistics analysis template for other developmental scientists by extracting low-level and high-level image features to understand how they influence visual perceptual development. We are currently exploring ways to enable this image statistics toolbox to be available to other researchers.

Applied contributions

Finally, the thesis has several applied contributions to industry and society. First, as a new diagnostic paediatric test of CVD, *ColourSpot* addresses the limitations of existing paediatric CVD tests and has great potential for the mass-screening and diagnosis of CVD in young children. Following the authorisations and approvals needed (e.g., product marking with UK conformity assessment), the long-term goal is that *ColourSpot* could be a feasible test for large-scale screening of CVD in the UK when children start school. Beyond its use in school screening programmes, *ColourSpot* also enables colour vision screening to be self-administered at home with a parent or at the clinic with the optometrist. Second, the thesis also increases awareness about the impact of CVD in the classroom for older children and adolescents and has implications for informing teachers and parents about the difficulties children with CVD experience with school tasks. This highlights the need for relevant support for these children. During *ColourSpot*'s development, factsheets for teachers and parents on CVD and its impact were developed and these also contribute to raising CVD awareness. Third, the thesis also makes an applied contribution to the design

industry, by identifying principles that can guide optimal pattern design for infants and toddlers. Throughout the PhD, I have collaborated with the professional designers of Cosatto Ltd who part-funded the PhD and translated the findings of the thesis to help the designers to optimise their designs for infant and toddler's vision and perception, and to create patterns that capture the development of visual perception. For example, informing designers that designs with faces and odd-one-out features will be looked at longer by infants and toddlers, but that designs could also be improved by designing character faces face-front instead of face-side. Or that designs with animals will likely be more attention capturing for toddlers than infants because of top-down processing and semantic knowledge. These findings also have implications for the design of other children's products and media like television, animation, films and book publishing.

6.2 Future research

In addition to the theoretical, methodological and applied contributions that the thesis makes, the thesis also raises a number of additional questions for future research.

The development of *ColourSpot* in Paper 1 means that the impact of CVD on children can now be studied more easily, with remote iPad testing as shown in Paper 2. If recruitment rates can be boosted, then there is good potential for this

method to enable large and diverse samples to be tested in studies of CVD. The null effect in Paper 2, whereby CVD was found not to have an effect on self-reported wellbeing or educational engagement, was not a firm null effect, and further data collection is required to be more confident of this. A replication of Paper 2 with a larger sample would be useful, but this study could also be extended to add other interesting factors. For example, in Paper 2, all children in the CVD group (except one) were aware that they had CVD before taking part in the study and an extension could include a sample of CVD children unaware of their condition. This would enable a test of the hypothesis that diagnosis of CVD, and subsequent support, could be protective of the negative impact of CVD on quality of life. In addition, including a larger age range would enable us to investigate whether the negative impact of CVD is cumulative as one develops or whether one develops strategies to mitigate the negative impact later in life.

Another question for future research is whether the severity and type of CVD is important when considering the impact of CVD – for example, it is predicted that dichromats feel a greater effect of CVD on their daily life than anomalous trichromats (Steward & Cole, 1989). Diagnosing CVD using the anomaloscope as well as *ColourSpot* would enable us to test these predictions and help us learn whether accurate *ColourSpot* classification criteria that distinguish CVD type can also be identified. Other modifications of *ColourSpot* could be developed and evaluated in further studies. For example, it could be possible to

develop an early years version of *ColourSpot* that would be more appropriate for children under the age of 4, perhaps presenting one target at a time rather than three targets. An adult version of *ColourSpot* can easily be developed by removing animations and verbal prompts, and further testing of adults with both *ColourSpot* and the anomaloscope following the same protocol as Paper 1 would enable the classification criteria for CVD in adults to be confirmed.

Paper 2 established that children with CVD lack confidence, and perhaps ability, in certain school tasks that involve colour, in particular, naming colours, using colour in art and being confident in selecting the right colour, doing maths or English exercises when they are coloured and doing things in science or design technology that involve colour. More detailed investigation of this using focus groups as well as further surveys could enable the development of a paediatric CVD quality of life questionnaire that is comparable to the one that Barry et al. (2017) have developed for adults. Such a questionnaire would be a valuable tool for educators, parents and optometrists to identify what support a child with CVD may need to mitigate any negative impact of CVD on their wellbeing, educational performance or more broadly. The process of completing the questionnaire with the child would also raise awareness of the types of issues that are and are not a problem for those with CVD, leading to more tailored and effective support.

The current thesis did not find any evidence that infants preferentially look at abstract images that have natural colour distributions that align with the natural blue-yellow daylight axis. Paper 4 outlined a series of possible explanations for the lack of effect for infant visual preference when a prior study has found an effect for infant sensitivity for chromatic scene statistics on a discrimination task (Skelton et al., 2021). Further research using additional methods, such as visual evoked potentials (for e.g., Crognale et al., 1998) could provide further evidence as to whether infant colour vision is aligned with chromatic scene statistics. If so, then a cross-environmental study which compares infants raised in environments with different chromatic scene statistics could be useful for deciphering whether colour vision aligns to natural scene statistics on an innate evolutionary timescale or whether colour vision tunes into the chromatic scene statistics of the environment the visual system is exposed to during development.

More broadly, further research on the role of various types of natural scene statistics in visual development is needed. Paper 3 failed to find any contribution of image statistics to infant and toddler visual preference, other than the dominant hue. However, due to the use of stylised abstract designs the range of the image statistics measures was limited compared to the range that is present in natural scenes. Presenting infants with a large set of natural images to see whether they would show a visual preference for images with certain natural scene statistics may reveal effects not revealed in Paper 3. The image analyses

developed in Paper 3 can be applied in further research to explore the effect of natural scene statistics in infant vision and perception. Also, future research could use head-mounted colour-calibrated cameras to capture images of an infant's visual environment from the infant's perspective (Jayaraman & Smith, 2020; L. Smith et al., 2011, 2015). The natural image statistics of images captured from infants' head-mounted cameras could be analysed to measure the statistical regularities of an infant's visual environment and to identify whether scene statistics change with development (e.g., as the child starts to crawl and walk). If the natural scene statistics of an infant's environment do change with development, the next question is whether these changes are accompanied by changes in the sensitivity of the visual system to those statistics. If so, that would provide further evidence that the visual system is calibrated by the environment and individual experience and provide a developmental trajectory of the sensitivity to natural scene statistics.

Paper 3 of the thesis was able to identify some visual features, such as faces, odd-one-out elements and preference for charteuse-ish and achromatic dominant hues, that appear to make infants and toddlers look longer at abstract patterns. Additional research could be done to identify the visual features that attract longer looking by infants and young children, to develop designs that are optimised for this age group. Being able to produce optimal designs for infants' and toddler's immature visual and perceptual systems could contribute to improving child engagement with children's products and may improve learning

from children's media such as toys, books and television. The very large interdisciplinary field of aesthetics has identified many principles of visual aesthetics in adults, and it is an interesting and practical question for further research whether these principles are at play for early visual preferences as well. As discussed earlier, visual preference is not equivalent to an aesthetic response (C. Taylor et al., 2013), and further research should consider when an aesthetic response develops and what factors influence its development. Given the large scale attempt to understand aesthetics, it is surprising that only a few studies have investigated the development of visual art preferences and the origins of aesthetic preference (Cacchione et al., 2011; Göksun et al., 2014; Krentz & Earl, 2013). Future research combining vision science (which enables precise characterisation of the stimulus) and developmental science could make a valuable contribution.

6.3 Conclusion

This thesis has provided further understanding of colour perception in infancy and childhood. Perhaps the most important contribution of the thesis has been to develop and evaluate a paediatric test of CVD, *ColourSpot*, which has good potential to be used for mass screening of CVD at the start of education. In addition, the thesis identifies the types of school tasks that children with CVD self-report having difficulty with; provides a protocol for conducting colour-calibrated remote iPad studies; identifies that faces and odd-one-out visual features and a charteuse-ish and achromatic dominant hue all appear to garner

pattern preference in infants and toddlers; establishes that infant visual preference does not increase for abstract images with natural blue-yellow colour distributions; and develops image analyses suitable for further research on the role of natural scene statistics in visual development. These contributions will hopefully benefit a range of people, from optometrists, parents and teachers interested in diagnosing CVD, to developmental and vision scientists testing theories of perceptual development or colour perception, to designers considering how to design products that are optimised for infants and children.

References

Abramov, I., Hainline, L., Turkel, J., Lemerise, E., Smith, H., Gordon, J., & Petry, S. (1984). Rocket-ship psychophysics. Assessing visual functioning in young children. *Investigative Ophthalmology and Visual Science*, *25*(11), 1307–1315.

Açık, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental Changes in Natural Viewing Behavior: Bottom-Up and Top-Down Differences between Children, Young Adults and Older Adults. *Frontiers in Psychology*, *1*(207), 1–14. https://doi.org/10.3389/fpsyg.2010.00207

Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behavior and Development*, *10*(2), 143–150. https://doi.org/10.1016/0163-6383(87)90029-4

Adler, S. A., Inslicht, S., Rovee-Collier, C., & Gerhardstein, P. C. (1998). Perceptual asymmetry and memory retrieval in 3-month-old infants. *Infant Behavior & Development*, *21*(2), 253–272. https://doi.org/10.1016/S0163-6383(98)90005-4

Adler, S. A., & Orprecio, J. (2006). The eyes have it: visual pop-out in infants and adults. *Developmental Science*, *9*(2), 189–206. https://doi.org/10.1111/j.1467-7687.2006.00479.x

Aks, D. J., & Sprott, J. C. (1996). Quantifying Aesthetic Preference for Chaotic

Patterns. *Empirical Studies of the Arts*, *14*(1), 1–16. https://doi.org/10.2190/6V31-7M9R-T9L5-CDG9

Albany-Ward, K. (2005). What do you really know about colour blindness? *British Journal of School Nursing*, *10*(4), 123–124.

Albany-Ward, K. (2011). Why colour really does matter. *Prep School Reflecting the Best in the Prep & Junior School World*, 34–35.

Albers, A., Gegenfurtner, K., & Nascimento, S. (2020). An independent contribution of colour to the aesthetic preference for paintings. *Vision Research*, *177*, 109–117. https://doi.org/10.1016/j.visres.2020.08.005

Altmann, C. S., Brachmann, A., & Redies, C. (2021). Liking of art and the perception of color. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(4), 545–564. https://doi.org/10.1037/xhp0000771

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. https://doi.org/10.1371/journal.pone.0180310

Álvaro, L., Moreira, H., Lillo, J., & Franklin, A. (2015). Color preference in red–green dichromats. *Proceedings of the National Academy of Sciences*, *112*(30), 9316–9321. https://doi.org/10.1073/pnas.1502104112

Amir, O., Biederman, I., & Hayworth, K. J. (2011). The neural basis for shape preferences. *Vision Research*, *51*(20), 2198–2206. https://doi.org/10.1016/j.visres.2011.08.015

Amso, D., Haas, S., & Markant, J. (2014). An Eye Tracking Investigation of Developmental Change in Bottom-up Attention Orienting to Faces in Cluttered Natural Scenes. *PLOS ONE*, *9*(1), e85701. https://doi.org/10.1371/journal.pone.0085701

Aslam, M. M. (2006). Are You Selling the Right Colour? A Cross-cultural Review of Colour as a Marketing Cue. *Journal of Marketing Communications*, *12*(1), 15–30. https://doi.org/10.1080/13527260500247827

Atick, J. J., Li, Z., & Redlich, A. N. (1993). What does post-adaptation color appearance reveal about cortical color representation? *Vision Research*, *33*(1), 123–129. https://doi.org/10.1016/0042-6989(93)90065-5

Atkinson, J., & Braddick, O. (2012). Visual attention in the first years: typical development and developmental disorders. *Developmental Medicine & Child Neurology*, *54*(7), 589–595. https://doi.org/10.1111/j.1469-8749.2012.04294.x

Atkinson, J., Braddick, O., & Moar, K. (1977). Development of contrast sensitivity over the first 3 months of life in the human infant. *Vision Research*, *17*(9), 1037–1044. https://doi.org/10.1016/0042-6989(77)90007-4

Atowa, U., Wajuihian, S., & Hansraj, R. (2019). A review of paediatric vision screening protocols and guidelines. *International Journal of Ophthalmology*, *12*(7), 1194–1201. https://doi.org/10.18240/ijo.2019.07.22

Azizoğlu, S., Crewther, S., Şerefhan, F., Barutchu, A., Göker, S., & Junghans, B. (2017). Evidence for the need for vision screening of school children in Turkey. *BMC Ophthalmology*, *17*(1), 230. https://doi.org/10.1186/s12886-017-0618-9

Balas, B., Lammers, S., & Saville, A. (2020). Neural sensitivity to natural texture statistics changes during middle childhood. *Journal of Vision*, *20*(11), 180. https://doi.org/10.1167/jov.20.11.180

Balas, B., Saville, A., & Schmidt, J. (2018). Neural sensitivity to natural texture statistics in infancy. *Developmental Psychobiology*, *60*(7), 765–774. https://doi.org/10.1002/dev.21764

Banks, M. S., & Dannemiller, J. L. (1987). Infant visual psychophysics. In P. Salapatek & L. B. Cohen (Eds.), *Handbook of infant perception* (Vol. 1, pp. 115–184). Academic Press.

Barbur, J., & Rodriguez-Carmona, M. (2015). Color vision changes in normal aging. In A. J. Elliott, M. D. Fairchild, & A. Franklin (Eds.), *Handbook of Color Psychology* (pp. 180–196). Cambridge University Press. https://doi.org/10.1017/CBO9781107337930.009

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory Communication*, *1*(01).

Barry, J., Mollan, S., Burdon, M., Jenkins, M., & Denniston, A. (2017). Development and validation of a questionnaire assessing the quality of life impact of Colour Blindness (CBQoL). *BMC Ophthalmology*, *17*(1), 1–7.

https://doi.org/10.1186/s12886-017-0579-z

Berman, M. G., Hout, M. C., Kardan, O., Hunter, M. R., Yourganov, G., Henderson, J. M., Hanayik, T., Karimi, H., & Jonides, J. (2014). The Perception of Naturalness Correlates with Low-Level Visual Features of Environmental Scenes. *PLOS ONE*, *9*(12), e114572. https://doi.org/10.1371/journal.pone.0114572

Birch, J. (1997). Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic and Physiological Optics*, *17*(5), 403–408. https://doi.org/10.1046/j.1475-1313.1997.97000227.x

Birch, J. (2001). *Diagnosis of defective colour vision*. Butterworth-Heinemann Ltd.

Birch, J. (2012). Worldwide prevalence of red-green color deficiency. *Journal of the Optical Society of America A*, *29*(3), 313–320. https://doi.org/10.1364/josaa.29.000313

Birch, J., & McKeever, L. (1993). Survey of the accuracy of new pseudoisochromatic plates. *Ophthalmic & Physiological Optics : The Journal of the British College of Ophthalmic Opticians (Optometrists)*, *13*(1), 35–40. https://doi.org/10.1111/j.1475-1313.1993.tb00423.x

Birch, J., & Platts, C. (1993). Colour vision screening in children: an evaluation of three pseudoisochromatic tests. *Ophthalmic and Physiological Optics*, *13*(4), 344–349. https://doi.org/10.1111/j.1475-1313.1993.tb00489.x

Blakemore, C., & Cooper, G. (1970). Development of the Brain depends on the

Visual Environment. *Nature*, *228*(5270), 477–478. https://doi.org/10.1038/228477a0

Block, S., Lee, D., Hoeppner, J., & Birr, A. (2004). Comparison of the Neitz Test of Color Vision to the Ishihara Color Vision tests and the anomaloscopic classification. *Investigative Ophthalmology & Visual Science*, *45*(13), 1384.

Bodduluri, L., Boon, M., & Dain, S. (2017). Evaluation of tablet computers for visual function assessment. *Behavior Research Methods*, *49*(2), 548–558. https://doi.org/10.3758/s13428-016-0725-1

Bodduluri, L., Boon, M., Ryan, M., & Dain, S. (2017). Impact of Gamification of Vision Tests on the User Experience. *Games for Health Journal*, *6*(4), 229–236. https://doi.org/10.1089/g4h.2016.0100

Bornstein, M. H. (1975). Qualities of color vision in infancy. *Journal of Experimental Child Psychology*, *19*(3), 401–419. https://doi.org/10.1016/0022-0965(75)90070-3

Bornstein, M. H. (1976). Infants' recognition memory for hue. *Developmental Psychology*, *12*(3), 185–191. https://doi.org/10.1037/0012-1649.12.3.185

Bornstein, M. H. (1978). Chromatic vision in infancy. *Advances in Child Development and Behavior*, *12*, 117–182. https://doi.org/10.1016/S0065-2407(08)60037-3

Bornstein, M. H., Arterberry, M. E., & Mash, C. (2011). Perceptual development. In M. H. Bornstein & M. E. Lambs (Eds.), *Developmental science: An advanced textbook* (pp. 303–351). Psychology Press.

Bornstein, M. H., Ferdinandsen, K., & Gross, C. G. (1981). Perception of symmetry in infancy. *Developmental Psychology*, *17*(1), 82–86. https://doi.org/10.1037/0012-1649.17.1.82

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. https://doi.org/10.1037/0096-1523.2.1.115

Bornstein, M. H., & Krinsky, S. J. (1985). Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes. *Journal of Experimental Child Psychology*, *39*(1), 1–19. https://doi.org/10.1016/0022-0965(85)90026-8

Bosten, J. (2019). The known unknowns of anomalous trichromacy. *Current Opinion in Behavioral Sciences*, *30*, 228–237. https://doi.org/10.1016/j.cobeha.2019.10.015

Bosten, J., Beer, R., & MacLeod, D. (2015). What is white? *Journal of Vision*, *15*(16), 5. https://doi.org/10.1167/15.16.5

Bosten, J., Mollon, J., Peterzell, D., & Webster, M. (2017). Individual differences as a window into the structure and function of the visual system. *Vision Research*, *141*, 1–3. https://doi.org/10.1016/j.visres.2017.11.003

Brennan, W. M., Ames, E. W., & Moore, R. W. (1966). Age Differences in Infants' Attention to Patterns of Different Complexities. *Science*, *151*(3708), 354–356. https://doi.org/10.1126/science.151.3708.354

Brettel, H., Viénot, F., & Mollon, J. D. (1997). Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America A*, *14*(10), 2647–2655. https://doi.org/10.1364/josaa.14.002647

Brown, A. M., & Lindsey, D. T. (2013). Infant color vision and color preferences: a tribute to Davida Teller. *Visual Neuroscience*, *30*(5–6), 243–250. https://doi.org/10.1017/S0952523813000114

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. https://doi.org/10.1098/rspb.1983.0090

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132. https://doi.org/10.1016/j.cognition.2011.06.010

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception & Psychophysics*, *43*(5), 494–507. https://doi.org/10.3758/BF03207885

Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, *26*(1), 157. https://doi.org/10.1364/ao.26.000157

Cacchione, T., Möhring, W., & Bertin, E. (2011). What is it about Picasso? Infants' categorical and discriminatory abilities in the visual arts. In *Psychology of Aesthetics, Creativity, and the Arts* (Vol. 5, Issue 4, pp. 370–

378). Educational Publishing Foundation. https://doi.org/10.1037/a0024129

Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-8*(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*, *15*(6), 379–383. https://doi.org/10.1111/j.0956-7976.2004.00688.x

Catherwood, D., Crassini, B., & Freiberg, K. (1987). The nature of infant memory for hue. *British Journal of Developmental Psychology*, *5*(4), 385–394. https://doi.org/10.1111/j.2044-835X.1987.tb01074.x

Catherwood, D., Crassini, B., & Freiberg, K. (1990a). The Relative Encoding of Colour and Shape by 5-Month-Old Infants. *The Australian Educational and Developmental Psychologist*, *7*(2), 1–3. https://doi.org/DOI: 10.1017/S0816512200026110

Catherwood, D., Crassini, B., & Freiberg, K. (1990b). The course of infant memory for hue. *Australian Journal of Psychology*, *42*(3), 277–285. https://doi.org/10.1080/00049539008260125

Catherwood, D., Skoien, P., & Holt, C. (1996). Colour pop-out in infant response to visual arrays. *British Journal of Developmental Psychology*, *14*(3), 315–326. https://doi.org/10.1111/j.2044-835X.1996.tb00708.x

Chakrabarti, S. (2018). Psychosocial aspects of colour vision deficiency: Implications for a career in medicine. *The National Medical Journal of India*,

31(2), 86. https://doi.org/10.4103/0970-258X.253167

Chan, X., Goh, S., & Tan, N. (2014). Subjects with colour vision deficiency in the community: What do primary care physicians need to know? *Asia Pacific Family Medicine*, *13*(1), 1–10. https://doi.org/10.1186/s12930-014-0010-3

Chia, A., Gazzard, G., Tong, L., Zhang, X., Sim, E. L., Fong, A., & Saw, S. M. (2008). Red-green colour blindness in Singaporean children. *Clinical and Experimental Ophthalmology*, *36*(5), 464–467. https://doi.org/10.1111/j.1442-9071.2008.01799.x

Choi, S. Y., & Hwang, J. M. (2009). Ishihara test in 3- to 6-year-old children. *Japanese Journal of Ophthalmology*, *53*(5), 455–457. https://doi.org/10.1007/s10384-009-0716-1

CIE. (1977). CIE Recommendations on Uniform Color Spaces, Color-Difference Equations, and Metric Color Terms. *Color Research & Application*, *2*(1), 5–6. https://doi.org/10.1002/j.1520-6378.1977.tb00102.x

Ciner, E., Dobson, V., Schmidt, P., Allen, D., Cyert, L., Maguire, M., Moore, B., Orel-Bixler, D., & Schultz, J. (1999). A survey of vision screening policy of preschool children in the United States. *Survey of Ophthalmology*, *43*(5), 445–457. https://doi.org/10.1016/s0039-6257(99)00021-1

Cohen, L. (1972). Attention-getting and attention-holding processes of infant visual preferences. *Child Development*, 869–879.

Cole, B. (1972). The Handicap of Abnormal Colour Vision. *The Australian*

Journal of Optometry, *55*(8), 304–310. https://doi.org/10.1111/j.1444-0938.1972.tb06271.x

Cole, B. (2004). The handicap of abnormal colour vision. *Clinical and Experimental Optometry*, *87*(4-5), 258–275. https://doi.org/10.1111/j.1444-0938.2004.tb05056.x

Cole, B. (2007). Assessment of inherited colour vision defects in clinical practice. *Clinical and Experimental Optometry*, *90*(3), 157–175. https://doi.org/10.1111/j.1444-0938.2007.00135.x

Cole, B. (2015). Re: Is screening for congenital colour vision deficiency in school students worthwhile? *Clinical and Experimental Optometry*, *98*(2), 193. https://doi.org/10.1111/cxo.12229

Cole, B., Lian, K., & Lakkis, C. (2006). The new Richmond HRR pseudoisochromatic test for colour vision is better than the Ishihara test. *Clinical and Experimental Optometry*, *89*(2), 73–80. https://doi.org/10.1111/j.1444-0938.2006.00015.x

Cole, G., & Wilkins, A. (2013). Fear of Holes. *Psychological Science*, *24*(10), 1980–1985. https://doi.org/10.1177/0956797613484937

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*, 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Colombo, J., Ryther, J. S., Frick, J. E., & Gifford, J. J. (1995). Visual pop-out in infants: Evidence for preattentive search in 3- and 4-month-olds.

Psychonomic Bulletin & Review, *2*(2), 266–268. https://doi.org/10.3758/BF03210968

Conway, B. R. (2009). Color Vision, Cones, and Color-Coding in the Cortex. *The Neuroscientist*, *15*(3), 274–290. https://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. https://doi.org/10.1523/JNEUROSCI.4348-10.2010

Courage, M. L., & Adams, R. J. (1990). Visual acuity assessment from birth to three years using the acuity card procedure: cross-sectional and longitudinal samples. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *67*(9), 713–718. https://doi.org/10.1097/00006324-199009000-00011

Cranwell, M. B., Pearce, B., Loveridge, C., & Hurlbert, A. C. (2015). Performance on the Farnsworth-Munsell 100-hue test is significantly related to nonverbal IQ. *Investigative Ophthalmology and Visual Science*, *56*(5), 3171–3178. https://doi.org/10.1167/iovs.14-16094

Crognale, M. A., Kelly, J. P., Weiss, A. H., & Teller, D. (1998). Development of the spatio-chromatic visual evoked potential (VEP): a longitudinal study. *Vision Research*, *38*(21), 3283–3292. https://doi.org/10.1016/S0042-6989(98)00074-1

Cubit, L. S., Canale, R., Handsman, R., Kidd, C., & Bennetto, L. (2021). Visual Attention Preference for Intermediate Predictability in Young Children. *Child Development*, *92*(2), 691–703. https://doi.org/10.1111/cdev.13536

Cumberland, P., Rahi, J. S., & Peckham, C. S. (2004). Impact of congenital colour vision deficiency on education and unintentional injuries: Findings from the 1958 British birth cohort. *British Medical Journal*, *329*(7474), 1074–1075. https://doi.org/10.1136/bmj.38176.685208.F7

Cumberland, P., Rahi, J. S., & Peckham, C. S. (2005). Impact of congenital colour vision defects on occupation. *Archives of Disease in Childhood*, *90*(9), 906–908. https://doi.org/10.1136/adc.2004.062067

Dain, S. (2010). Evaluation of “Colour Vision Testing Made Easy.” In J. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and Defective Colour Vision*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525301.003.0035

Dain, S., & Almerdef, A. (2016). Colorimetric evaluation of iPhone apps for colour vision tests based on the Ishihara test. *Clinical and Experimental Optometry*, *99*(3), 264–273. https://doi.org/10.1111/cxo.12370

Dain, S., & Ling, B. (2009). Cognitive abilities of children on a gray seriation test. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *86*(6), E701-707. https://doi.org/10.1097/OPX.0b013e3181a59d46

Damon, F., Mottier, H., Méary, D., & Pascalis, O. (2017). A Review of

Attractiveness Preferences in Infancy: From Faces to Objects. *Adaptive Human Behavior and Physiology*, *3*(4), 321–336. https://doi.org/10.1007/s40750-017-0071-2

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

Dannemiller, J., & Stephens, B. (1988). A critical test of infant pattern preference models. *Child Development*, *59*(1), 210–216.

Davies, I., & Franklin, A. (2002). Categorical similarity may affect colour pop-out in infants after all. *British Journal of Developmental Psychology*, *20*(2), 185–203. https://doi.org/10.1348/026151002166406

de Fez, D., Luque, M., García-Domene, M., Caballero, M., & Camps, V. (2018). Can Applications Designed to Evaluate Visual Function Be Used in Different iPads? *Optometry and Vision Science*, *95*(11), 1054–1063. https://doi.org/10.1097/OPX.0000000000001293

de Fez, D., Luque, M., Garcia-Domene, M., Camps, V., & Pineero, D. (2016). Colorimetric characterization of mobile devices for vision applications. *Optometry and Vision Science*, *93*(1), 85–93. https://doi.org/10.1097/OPX.0000000000000752

de Fez, D., Luque, M., Matea, L., Piñero, D., & Camps, V. (2018). New iPAD-based test for the detection of color vision deficiencies. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *256*(12), 2349–2360.

https://doi.org/10.1007/s00417-018-4154-y

De Valois, K., & De Valois, R. (2000). Color Vision. In K. De Valois (Ed.), *Seeing* (pp. 129–175). Elsevier. https://doi.org/10.1016/B978-012443760-9/50006-4

De Valois, R., Abramov, I., & Jacobs, G. (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America*, *56*(7), 966–977. https://doi.org/10.1364/JOSA.56.000966

Dekker, T. M., Farahbakhsh, M., Atkinson, J., Braddick, O. J., & Jones, P. R. (2020). Development of the spatial contrast sensitivity function (CSF) during childhood: Analysis of previous findings and new psychophysical data. *Journal of Vision*, *20*(13), 4. https://doi.org/10.1167/jov.20.13.4

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

Department of Health. (2009). *Healthy Child Programme: From 5 to 19 years old*.

Derrington, A., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*, 241–265. https://doi.org/10.1113/jphysiol.1984.sp015499

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Diez, M., Luque, M., Capilla, P., Gomez, J., & de Fez, D. (2001). Detection and

assessment of color vision anomalies and deficiencies in children. *Journal of Pediatric Ophthalmology and Strabismus*, *38*(4), 191–195. https://doi.org/10.3928/0191-3913-20010701-05

Dzulkifli, M., & Mustafar, M. (2013). The influence of colour on memory performance: a review. *The Malaysian Journal of Medical Sciences : MJMS*, *20*(2), 3–9.

Ellemberg, D., Hansen, B. C., & Johnson, A. (2012). The developing visual system is not optimally sensitive to the spatial statistics of natural images. *Vision Research*, *67*, 1–7. https://doi.org/10.1016/j.visres.2012.06.018

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of Experimental Psychology. General*, *136*(1), 154–168. https://doi.org/10.1037/0096-3445.136.1.154

Elliot, A. J., & Niesta, D. (2008). Romantic red: Red enhances men's attraction to women. *Journal of Personality and Social Psychology*, *95*(5), 1150–1164. https://doi.org/10.1037/0022-3514.95.5.1150

Elliot, A. J., Niesta Kayser, D., Greitemeyer, T., Lichtenfeld, S., Gramzow, R. H., Maier, M. A., & Liu, H. (2010). Red, rank, and romance in women viewing men. *Journal of Experimental Psychology: General*, *139*(3), 399–417. https://doi.org/10.1037/a0019689

Emberson, L. L. (2017). How Does Experience Shape Early Development? Considering the Role of Top-Down Mechanisms. *Advances in Child*

Development and Behavior., *52*, 1–41. https://doi.org/10.1016/bs.acdb.2016.10.001

Erhart, M., Ottova, V., Gaspar, T., Jericek, H., Schnohr, C., Alikasifoglu, M., Morgan, A., & Ravens-Sieberer, U. (2009). Measuring mental health and well-being of school-children in 15 European countries using the KIDSCREEN-10 Index. *International Journal of Public Health*, *54 Suppl 2*, 160–166. https://doi.org/10.1007/s00038-009-5407-7

Fanlo Zarazaga, A., Gutiérrez Vásquez, J., & Pueyo Royo, V. (2019). Review of the main colour vision clinical assessment tests. *Archivos de La Sociedad Espanola de Oftalmologia*, *94*(1), 25–32. https://doi.org/10.1016/j.oftal.2018.08.006

Fantz, R. L. (1958). Pattern Vision in Young Infants. *The Psychological Record*, *8*, 43–47.

Fantz, R. L. (1961). The Origin of Form Perception. *Scientific American*, *204*(5), 66–73. https://doi.org/10.1038/scientificamerican0561-66

Fantz, R. L. (1963). Pattern Vision in Newborn Infants. *Science*, *140*(3564), 296–297. https://doi.org/10.1126/science.140.3564.296

Fantz, R. L. (1965). Visual perception from birth as shown by pattern selectivity. *Annals of the New York Academy of Sciences*, *118*(21), 793–814. https://doi.org/10.1111/j.1749-6632.1965.tb40152.x

Fantz, R. L., & Fagan, J. F. 3rd. (1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child*

Development, *46*(1), 3–18.

Fantz, R. L., & Miranda, S. (1975). Newborn Infant Attention to Form of Contour. *Child Development*, *46*(1), 224. https://doi.org/10.2307/1128853

Fantz, R. L., & Nevis, S. (1967). Pattern preferences and perceptual-cognitive development in early infancy. *Merrill-Palmer Quarterly of Behavior and Development*, *13*(1), 77–108.

Farnsworth, D. (1943). The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *Journal of the Optical Society of America*, *33*(10), 568–578. https://doi.org/10.1364/josa.33.000568

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: effects of contrast polarity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(47), 17245–17250. https://doi.org/10.1073/pnas.0502205102

Fernandez, D., & Wilkins, A. J. (2008). Uncomfortable images in art and nature. *Perception*, *37*(7), 1098–1113. https://doi.org/10.1068/p5814

Field, D. (1994). What is the goal of sensory coding? *Neural Computation*, *6*(4), 559–601. https://doi.org/10.1162/neco.1994.6.4.559

Field, D., & Brady, N. (1997). Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, *37*(23), 3367–3383. https://doi.org/10.1016/s0042-6989(97)00181-8

Field, T., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face

discrimination by the newborn. *Infant Behavior and Development*, *7*(1), 19–25. https://doi.org/10.1016/S0163-6383(84)80019-3

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15822–15826. https://doi.org/10.1073/pnas.232472899

Fish, A. L., Alketbi, M., & Baillif, S. (2020). Evaluation of a New Test for the Diagnosis of Congenital Dyschromatopsia in Children: the Color Vision Evaluation Test. *American Journal of Ophthalmology*, *223*, 348–358. https://doi.org/10.1016/j.ajo.2020.11.003

Fisher, C. B., Ferdinandsen, K., & Bornstein, M. H. (1981). The role of symmetry in infant form discrimination. *Child Development*, 457–462.

Fletcher, R. (1984). The City University Colour Vision Test. *Ophthalmic and Physiological Optics*, *4*(3), 279–279. https://doi.org/10.1111/j.1475-1313.1984.tb00367.x

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

Franchak, J. M. (2020). Visual exploratory behavior and its development. In K. Federmeier & E. Schotter (Eds.), *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 73, pp. 59–94). Academic Press. https://doi.org/10.1016/bs.plm.2020.07.001

Franchak, J. M., Heeger, D. J., Hasson, U., & Adolph, K. E. (2016). Free

Viewing Gaze Behavior in Infants and Adults. *Infancy*, *21*(3), 262–287. https://doi.org/10.1111/infa.12119

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*, 13–26. https://doi.org/10.1016/j.jecp.2013.08.012

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170. https://doi.org/10.1016/j.cognition.2008.11.010

Franklin, A., Bevis, L., Ling, Y., & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science*, *13*(2), 346–354. https://doi.org/10.1111/j.1467-7687.2009.00884.x

Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology*, *22*(3), 349–377. https://doi.org/10.1348/0261510041552738

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. https://doi.org/10.1348/026151007X256672

Gegenfurtner, K. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, *4*(7), 563–572. https://doi.org/10.1038/nrn1138

Gegenfurtner, K., & Kiper, D. (2003). Colour vision. *Annual Review of Neuroscience*, *26*(1), 181–206.

https://doi.org/10.1146/annurev.neuro.26.041002.131116

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, *59*, 167–192. https://doi.org/10.1146/annurev.psych.58.110405.085632

Gerhardstein, P., Renner, P., & Rovee-Collier, C. (1999). The roles of perceptual and categorical similarity in colour pop-out in infants. *British Journal of Developmental Psychology*, *17*(3), 403–420. https://doi.org/10.1348/026151099165366

Gerhardstein, P., & Rovee-Collier, C. (2002). The Development of Visual Search in Infants and Very Young Children. *Journal of Experimental Child Psychology*, *81*(2), 194–215. https://doi.org/10.1006/jecp.2001.2649

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces Attract Infants' Attention in Complex Displays. *Infancy*, *14*(5), 550–562. https://doi.org/10.1080/15250000903144199

Göksun, T., Kranjec, A., & Chatterjee, A. (2014). The development of visual art preferences. *Proceedings of the 23rd International Association of Empirical Aesthetics*, 1–4.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Goulart, P. R. K., Bandeira, M. L., Tsubota, D., Oiwa, N. N., Costa, M. F., & Ventura, D. F. (2008). A computer-controlled color vision test for children based on the Cambridge Colour Test. *Visual Neuroscience*, *25*(3), 445–450. https://doi.org/10.1017/S0952523808080589

Graham, D., & Field, D. (2007). Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision*, *21*(1–2), 149–164. https://doi.org/10.1163/156856808782713771

Graham, D., & Redies, C. (2010). Statistical regularities in art: Relations with visual coding and perception. *Vision Research*, *50*(16), 1503–1509. https://doi.org/10.1016/j.visres.2010.05.002

Granzier, J. J. M., & Valsecchi, M. (2014). Variations in daylight as a contextual cue for estimating season, time of day, and weather conditions. *Journal of Vision*, *14*(1), 22. https://doi.org/10.1167/14.1.22

Grassivaro Gallo, P., Oliva, S., Lantieri, P., & Viviani, F. (2002). Colour Blindness in Italian Art High School Students. *Perceptual and Motor Skills*, *95*(3), 830–834. https://doi.org/10.2466/pms.2002.95.3.830

Grassivaro Gallo, P., Panza, M., Viviani, F., & Lantieri, P. (1998). Congenital dyschromatopsia and school achievement. *Perceptual and Motor Skills*, *86*(2), 563–569. https://doi.org/10.2466/pms.1998.86.2.563

Greenberg, D. J., & Blue, S. Z. (1975). Visual Complexity in Infancy: Contour or Numerosity? *Child Development*, *46*(2), 357–363. https://doi.org/10.2307/1128128

Hansen, B. C., & Hess, R. F. (2006). Discrimination of amplitude spectrum slope in the fovea and parafovea and the local amplitude distributions of natural scene imagery. *Journal of Vision*, *6*(7), 3. https://doi.org/10.1167/6.7.3

Hardy, L. H., Rand, G., & Rittler, M. C. (1954). H–R–R Polychromatic Plates. *Journal of the Optical Society of America*, *44*(7), 509–523. https://doi.org/10.1364/josa.44.000509

Harrington, S., Davison, P. A., & O'dwyer, V. (2021). Prevalence of colour vision deficiency in the Republic of Ireland schoolchildren and associated socio-demographic factors. *Clinical and Experimental Optometry*, *104*(1), 48–55. https://doi.org/10.1111/cxo.13072

Haskett, M. K., & Hovis, J. K. (1987). Comparison of the standard pseudoisochromatic plates to the Ishihara color vision test. *Optometry and Vision Science*, *64*(3), 211–216. https://doi.org/10.1097/00006324-198703000-00008

Helo, A., van Ommen, S., Pannasch, S., Danteny-Dordoigne, L., & Rämä, P. (2017). Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers. *Infant Behavior and Development*, *49*, 248–266. https://doi.org/10.1016/j.infbeh.2017.09.008

Hendrickson, A. E., & Yuodelis, C. (1984). The Morphological Development of the Human Fovea. *Ophthalmology*, *91*(6), 603–612. https://doi.org/10.1016/S0161-6420(84)34247-6

Hering, E. (1874). *Outlines of a Theory of the Light Sense*. Harvard University Press.

Hibbard, P. B., & O'Hare, L. (2015). Uncomfortable images produce non-sparse responses in a model of primary visual cortex. *Royal Society Open Science*, *2*(2), 140535. https://doi.org/10.1098/rsos.140535

Hilbert, D. R. (1992). What Is Color Vision? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *68*(3), 351–370. http://www.jstor.org/stable/4320360

Holroyd, E., & Hall, D. (1997). A re-appraisal of screening for colour vision impairments. *Child: Care, Health and Development*, *23*(5), 391–398. https://doi.org/10.1111/j.1365-2214.1997.tb00906.x

Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development*, *16*(4), 405–422. https://doi.org/10.1016/0163-6383(93)80001-O

Hopkins, S., Sampson, G., Hendicott, P., & Wood, J. (2013). Review of guidelines for children's vision screenings. *Clinical and Experimental Optometry*, *96*(5), 443–449. https://doi.org/10.1111/cxo.12029

Hovis, J., Leat, S., Heffernan, S., & Epp, K. (2002). The validity of the University of Waterloo Colored Dot Test for color vision testing in adults and preschool children. *Optometry and Vision Science*, *79*(4), 241–253. https://doi.org/10.1097/00006324-200204000-00011

Hubel, D. H., & Wiesel, T. N. (2004). *Brain and visual perception: the story of a*

25-year collaboration. Oxford University Press.

Hunter, R., & Pointer, M. (2011). *Measuring colour*. John Wiley & Sons, Ltd.

Hurlbert, A. (2007). Colour constancy. *Current Biology*, *17*(21), R906–R907. https://doi.org/10.1016/j.cub.2007.08.022

Hurlbert, A., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, *17*(16), R623–R625. https://doi.org/10.1016/j.cub.2007.06.022

Ibarra, F. F., Kardan, O., Hunter, M. R., Kotabe, H. P., Meyer, F. A. C., & Berman, M. G. (2017). Image feature types and their predictions of aesthetic preference and naturalness. *Frontiers in Psychology*, *8*, 632.

Ichikawa, H., Hukami, K., Tanabe, S., & Kawakami, G. (1979). Standard Pseudoisochromatic Plates Part 2. *Clinical and Experimental Optometry*, *62*, 362. https://doi.org/10.1111/j.1444-0938.1979.tb02331.x

Isherwood, Z., Clifford, C., Schira, M., Roberts, M., & Spehar, B. (2021). Nice and slow: Measuring sensitivity and visual preference toward naturalistic stimuli varying in their amplitude spectra in space and time. *Vision Research*, *181*, 47–60. https://doi.org/10.1016/j.visres.2021.01.001

Ishihara, S. (1917). *Tests for color-blindness*. Handaya.

Ishihara, S. (1974). *Tests for Colour-Blindness*. Kanehara Shuppan.

Ishihara, S. (1989). *The series of plates designed as a test for colour-deficiency*. Kanehara.

Ishihara, S. (1998). *Ishihara's tests for colour deficiency*. Kanehara.

Ishihara, S. (2002). *Ishihara's tests for colour deficiency*. Kanehara.

Ishihara, S., & Ishihara, M. (1943). *Ishihara's Design Charts for Colour Deficiency of Unlettered Persons*. Isshinkai Foundation.

Ishihara, S., & Ishihara, M. (2016). *Ishihara's design charts for colour deficiency of Unlettered Persons*. Kanehara Trading Inc.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Jadhav, A., Kumar Sg, P., & Kundu, S. (2017). Importance of colour vision testing in school based eye health examination. *Community Eye Health*, *30*(98), S24–S25.

Jadva, V., Hines, M., & Golombok, S. (2010). Infants' Preferences for Toys, Colors, and Shapes: Sex Differences and Similarities. *Archives of Sexual Behavior*, *39*(6), 1261–1273. https://doi.org/10.1007/s10508-010-9618-z

James, W. (1890). *The Principles of Psychology* (Vol. 1). Dover Publications.

Jayaraman, S., & Smith, L. B. (2020). The infant's visual world: Everyday statistics for visual learning. In J. J. Lockman & S. Tamis-LeMonda, C. (Eds.), *The Cambridge Handbook of Infant Development: Brain, Behavior, and Cultural Context* (pp. 553–579). Cambridge University Press.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*(10), 766–774. https://doi.org/10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Judd, D. (1945). Standard Response Functions for Protanopic and Deuteranopic Vision. *Journal of the Optical Society of America*, *35*(3), 199–221. https://doi.org/10.1364/JOSA.35.000199

Jurasevska, K., Ozolinsh, M., Fomins, S., Gutmane, A., Zutere, B., Pausus, A., & Karitans, V. (2014). Color-discrimination threshold determination using pseudoisochromatic test plates. *Frontiers in Psychology*, *5*, 1–7. https://doi.org/10.3389/fpsyg.2014.01376

Juricevic, I., Land, L., Wilkins, A., & Webster, M. A. (2010). Visual discomfort and natural image statistics. *Perception*, *39*(7), 884–899. https://doi.org/10.1068/p6656

Kadooka, K., & Franchak, J. M. (2020). Developmental changes in infants' and children's attention to faces and salient regions vary across and within video stimuli. *Developmental Psychology*, *56*(11), 2065–2079. https://doi.org/10.1037/dev0001073

Kaplan, E., Lee, B. B., & Shapley, R. M. (1990). New views of primate retinal function. *Progress in Retinal Research*, *9*, 273–336.

Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., Yourganov, G., Jonides, J., & Berman, M. G. (2015). Is the preference of natural versus man-made scenes driven by bottom-up processing of the

visual features of nature? *Frontiers in Psychology*, *6*, 471. https://doi.org/10.3389/fpsyg.2015.00471

Kauppinen-Räisänen, H., & Luomala, H. T. (2010). Exploring consumers' product-specific colour meanings. *Qualitative Market Research: An International Journal*, *13*(3), 287–308. https://doi.org/10.1108/13522751011053644

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLoS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Kimura, A., Wada, Y., Yang, J., Otsuka, Y., Dan, I., Masuda, T., Kanazawa, S., & Yamaguchi, M. K. (2010). Infants' recognition of objects using canonical color. *Journal of Experimental Child Psychology*, *105*(3), 256–263. https://doi.org/10.1016/j.jecp.2009.11.002

Kinnear, P. R., & Sahraie, A. (2002). New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5-22 and for age decades 30-70. *British Journal of Ophthalmology*, *86*(12), 1408–1411. https://doi.org/10.1136/bjo.86.12.1408

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Knoblauch, K., Vital-Durand, F., & Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research*, *41*(1), 23–36. https://doi.org/10.1016/S0042-6989(00)00205-4

Kovács, I. (2000). Human development of perceptual organization. *Vision Research*, *40*(10), 1301–1310. https://doi.org/10.1016/S0042-6989(00)00055-9

Kránicz, B. (2016). Daylight Illuminants. In M. R. Luo (Ed.), *Encyclopedia of Color Science and Technology* (pp. 553–558). Springer New York. https://doi.org/10.1007/978-1-4419-8071-7_320

Krauskopf, J., & Gegenfurtner, K. (1992). Color discrimination and adaptation. *Vision Research*, *32*(11), 2165–2175. https://doi.org/10.1016/0042-6989(92)90077-V

Krauskopf, J., Williams, D., & Heeley, D. (1982). Cardinal directions of color space. *Vision Research*, *22*(9), 1123–1131. https://doi.org/10.1016/0042-6989(82)90077-3

Krentz, U. C., & Earl, R. K. (2013). The baby as beholder: Adults and infants have common preferences for original art. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(2), 181–190. https://doi.org/10.1037/a0030691

Kulke, L., Atkinson, J., & Braddick, O. (2017). Neural mechanisms of attention become more specialised during infancy: Insights from combined eye tracking and EEG. *Developmental Psychobiology*, *59*(2), 250–260. https://doi.org/10.1002/dev.21494

Kwon, M. K., Setoodehnia, M., Baek, J., Luck, S., & Oakes, L. (2016). The development of visual search in infancy: Attention to faces versus salience. *Developmental Psychology*, *52*(4), 537–555. https://doi.org/10.1037/dev0000080

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. https://doi.org/10.1016/j.visres.2007.03.011

Lakowski, R. (1969). Theory and practice of colour vision testing: A review. Part 2. *Occupational and Environmental Medicine*, *26*(4), 265–288. https://doi.org/10.1136/oem.26.4.265

Lampe, J. M., Doster, M. E., & Beal, B. B. (1973). Summary of a three-year study of academic and school achievement between color-deficient and normal primary age pupils: phase two. *Journal of School Health*, *43*(5), 309–311. https://doi.org/10.1111/j.1746-1561.1973.tb03718.x

Land, E. (1983). Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(16), 5163–5169. https://doi.org/10.1073/pnas.80.16.5163

Lanthony, P. (1978). The desaturated panel D-15. *Documenta Ophthalmologica*, *46*(1), 185–189. https://doi.org/10.1007/BF00174107

Larsson, E. E. C., & von Stumm, S. (2015). Seeing red? The effect of colour on

intelligence test performance. *Intelligence*, *48*, 133–136. https://doi.org/10.1016/j.intell.2014.11.007

Lee, D., Cotter, S., & French, A. (1997). Evaluation of Kojima-Matsubara color vision test plates: validity in young children. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *74*(9), 726–731. https://doi.org/10.1097/00006324-199709000-00020

Lee, J., & Stromeyer, C. (1989). Contribution of human short-wave cones to luminance and motion detection. *The Journal of Physiology*, *413*, 563–593. https://doi.org/10.1113/jphysiol.1989.sp017669

Lennie, P., Pokorny, J., & Smith, V. (1993). Luminance. *Journal of the Optical Society of America A*, *10*(6), 1283–1293. https://doi.org/10.1364/JOSAA.10.001283

Lichtenfeld, S., Elliot, A. J., Maier, M. A., & Pekrun, R. (2012). Fertile Green. *Personality and Social Psychology Bulletin*, *38*(6), 784–797. https://doi.org/10.1177/0146167212436611

Lillo, J., Álvaro, L., & Moreira, H. (2014). An experimental method for the assessment of color simulation tools. *Journal of Vision*, *14*(8), 15. https://doi.org/10.1167/14.8.15

Lillo, J., Moreira, H., Álvaro, L., & Davies, I. (2014). Use of basic color terms by red–green dichromats: 1. General description. *Color Research & Application*, *39*(4), 360–371. https://doi.org/https://doi.org/10.1002/col.21803

Ling, B., & Dain, S. (2018). Development of color vision discrimination during childhood: differences between Blue–Yellow, Red–Green, and achromatic thresholds. *Journal of the Optical Society of America A*, *35*(4), B35–B42. https://doi.org/10.1364/josaa.35.000b35

Linksz, A. (1966). The Farnsworth panel D-15 test. *American Journal of Ophthalmology*, *62*(1), 27–37. https://doi.org/10.1016/0002-9394(66)91673-4

Lobue, V., & Deloache, J. S. (2011). Pretty in pink: The early development of gender-stereotyped colour preferences. *The British Journal of Developmental Psychology*, *29*, 656–667. https://doi.org/10.1111/j.2044-835X.2011.02027.x

Logan, N. S., & Gilmartin, B. (2004). School vision screening, ages 5-16 years: The evidence-base for content, provision and efficacy. *Ophthalmic and Physiological Optics*, *24*(6), 481–492. https://doi.org/10.1111/j.1475-1313.2004.00247.x

Long, J. A., Honson, V., Katalinic, P., & Dain, S. J. (2015). Re: Is screening for congenital colour vision deficiency in school students worthwhile? A review. *Clinical and Experimental Optometry*, *98*(2), 192. https://doi.org/10.1111/cxo.12246

MacLeod, D., & Boynton, R. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, *69*(8), 1183–1186. https://doi.org/10.1364/JOSA.69.001183

Mallon, B., Redies, C., & Hayn-Leichsenring, G. (2014). Beauty in abstract paintings: perceptual contrast and statistical properties. *Frontiers in Human Neuroscience*, *8*, 161.

Mäntyjärvi, M. (1987). An Evaluation of the Standard Pseudoisochromatic Plates (SPP 1) in Clinical Use. In G. Verriest (Ed.), *Colour Vision Deficiencies VIII. Documenta Ophthalmologica Proceedings Series* (pp. 125–131). Springer. https://doi.org/10.1007/978-94-009-4275-2_18

Mäntyjärvi, M. (1991a). Colour Vision Testing in Pre-School-Aged Children. *Ophthalmologica*, *202*(3), 147–151. https://doi.org/10.1159/000310196

Mäntyjärvi, M. (1991b). Velhagen Pflügertrident pseudoisochromatic plates in screening congenital red-green vision defects. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *229*(2), 145–146. https://doi.org/10.1007/bf00170546

Mäntyjärvi, M., Tenhunen, J., & Partanen, J. (2000). Screening color vision in preschool-aged children. *Color Research & Application*, *26*(S1), S250–S252. https://doi.org/10.1002/1520-6378(2001)26:1+<::AID-COL52>3.0.CO;2-I

Marechal, M., Delbarre, M., Tesson, J., Lacambre, C., Lefebvre, H., & Froussart-Maille, F. (2018). Color vision tests in pilots' medical assessments. *Aerospace Medicine and Human Performance*, *89*(8), 737–743. https://doi.org/10.3357/AMHP.5009.2018

Mashige, K. P. (2019). Impact of congenital color vision defect on color-related

tasks among schoolchildren in Durban, South Africa. *Clinical Optometry*, *11*, 97–102. https://doi.org/10.2147/OPTO.S204332

Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When art moves the eyes: A behavioral and eye-tracking study. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0037285

Mather, G. (2018). Visual Image Statistics in the History of Western Art. *Art & Perception*, *6*(2–3), 97–115. https://doi.org/https://doi.org/10.1163/22134913-20181092

Mather, G. (2020). Aesthetic Image Statistics Vary with Artistic Genre. *Vision*, *4*(1), 10. https://doi.org/10.3390/vision4010010

Matsubara, H., & Kojima, K. (1957). *Colour vision test plates for the infants*. Handaya.

Maurer, D., Lewis, T. L., Brent, H. P., & Levin, A. V. (1999). Rapid Improvement in the Acuity of Infants After Visual Input. *Science*, *286*(5437), 108 LP – 110. https://doi.org/10.1126/science.286.5437.108

Mayer, D. L., & Dobson, V. (1982). Visual acuity development in infants and young children, as assessed by operant preferential looking. *Vision Research*, *22*(9), 1141–1151. https://doi.org/10.1016/0042-6989(82)90079-7

McDermott, K. C., & Webster, M. A. (2012). Uniform color spaces and natural image statistics. *Journal of the Optical Society of America A*, *29*(2), A182.

https://doi.org/10.1364/josaa.29.00a182

McManus, I. C., Jones, A. L., & Cottrell, J. (1981). The Aesthetics of Colour. *Perception*, *10*(6), 651–666. https://doi.org/10.1068/p100651

Mehta, B., Sowden, P. T., & Grandison, A. (2018). Does deuteranomaly place children at a disadvantage in educational settings? In L. W. MacDonald, C. P. Biggam, & G. V. Paramei (Eds.), *Progress in Colour Studies. Cognition, language and beyond.* (pp. 341–355). John Benjamins Publishing Company. https://doi.org/10.1075/z.217.18meh

Meidenbauer, K. L., Stenfors, C. U. D., Young, J., Layden, E. A., Schertz, K. E., Kardan, O., Decety, J., & Berman, M. G. (2019). The gradual development of the preference for natural environments. *Journal of Environmental Psychology*, *65*, 101328. https://doi.org/10.1016/j.jenvp.2019.101328

Mollon, J., Astell, S., & Reffin, J. (1991). A minimalist test of colour vision. *Documenta Ophthalmologica Proceedings Series*, *54*, 59–67. https://doi.org/10.1007/978-94-011-3774-4_8

Montagner, C., Linhares, J. M. M., Vilarigues, M., & Nascimento, S. M. C. (2016). Statistics of colors in paintings and natural scenes. *Journal of the Optical Society of America A*, *33*(3), A170–A177. https://doi.org/10.1364/JOSAA.33.00A170

Moreira, H., Álvaro, L., Melnikova, A., & Lillo, J. (2018). Colorimetry and Dichromatic Vision. In *Colorimetry and Image Processing*. InTech. https://doi.org/10.5772/intechopen.71563

Mori, T., Endou, Y., & Nakayama, A. (1996). Fractal Analysis and Aesthetic Evaluation of Geometrically Overlapping Patterns. *Textile Research Journal*, *66*(9), 581–586. https://doi.org/10.1177/004051759606600906

Morrone, M. C., Burr, D. C., & Fiorentini, A. (1990). Development of Contrast Sensitivity and Acuity of the Infant Colour System. *Proceedings: Biological Sciences*, *242*(1304), 134–139.

Morrone, M. C., Burr, D. C., & Fiorentini, A. (1993). Development of infant contrast sensitivity to chromatic stimuli. *Vision Research*, *33*(17), 2535–2552. https://doi.org/10.1016/0042-6989(93)90133-H

Nagel, W. A. (1907). I. Zwei Apparate für die augenärztliche Funktionsprüfung. *Ophthalmologica*, *17*(3), 201–222. https://doi.org/10.1159/000291204

Nakauchi, S., Kondo, T., Higashi, H., Linhares, J., & Nascimento, S. (2018). Color statistics underlying preference judgement for art paintings. *Journal of Vision*, *18*(10), 867. https://doi.org/10.1167/18.10.867

Nascimento, S., Linhares, J., João, C., Amano, K., Alfaro, C., & Bailão, A. (2017). The colors of paintings and viewers' preferences. *Vision Research*, *130*, 76–84. https://doi.org/10.1016/j.visres.2016.11.006

Nascimento, S., Marit Albers, A., & Gegenfurtner, K. (2021). Naturalness and aesthetics of colors – Preference for color compositions perceived as natural. *Vision Research*, *185*, 98–110. https://doi.org/10.1016/j.visres.2021.03.010

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. https://doi.org/10.1126/science.2937147

Neitz, M., & Neitz, J. (1995). Numbers and ratios of visual pigment genes for normal red-green color vision. *Science*, *267*(5200), 1013–1016. https://doi.org/10.1126/science.7863325

Neitz, M., & Neitz, J. (2000). Molecular genetics of color vision and color vision defects. *Archives of Ophthalmology*, *118*(5), 691–700. https://doi.org/10.1001/archopht.118.5.691

Neitz, M., & Neitz, J. (2001). A new mass screening test for color-vision deficiencies in children. *Color Research and Application*, *26*(S1), S239–S249. https://doi.org/10.1002/1520-6378(2001)26:1+<::aid-col51>3.0.co;2-l

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*(1–2), 3–18. https://doi.org/https://doi.org/10.1002/icd.239

Nelson, C. A., Thomas, K. M., & De Haan, M. (2012). *Neuroscience of cognitive development: The role of experience and the developing brain*. John Wiley & Sons.

Nguyen, L., Do, E., Chia, A., Wang, Y., & Duh, H. (2014). DoDo game, a color vision deficiency screening test for young children. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2289–2292. https://doi.org/10.1145/2556288.2557334

Nguyen, L., Lu, W., Do, E., Chia, A., & Wang, Y. (2014). Using digital game as

clinical screening test to detect color deficiency in young children. *ACM International Conference Proceeding Series*, 337–340. https://doi.org/10.1145/2593968.2610486

Nguyen, L., & Spehar, B. (2021). Visual adaptation to natural scene statistics and visual preference. *Vision Research*, *180*, 87–95. https://doi.org/10.1016/j.visres.2020.11.011

Nithiyaananthan, P. J., Kaur, S., Anne, B. T., Eliana, A. N., & Mahadir, A. (2020). Behavioural Issues among Primary Schoolchildren with Colour Vision Deficiency. *IIUM Medical Journal Malaysia*, *19*(3), 37–45.

Norcia, A. M., Tyler, C. W., & Hamer, R. D. (1990). Development of contrast sensitivity in the human infant. *Vision Research*, *30*(10), 1475–1486. https://doi.org/10.1016/0042-6989(90)90028-J

O'Hare, L., & Hibbard, P. B. (2011). Spatial frequency and visual discomfort. *Vision Research*, *51*(15), 1767–1777. https://doi.org/10.1016/j.visres.2011.06.002

Oduntan, O. A., Mashige, K. P., & Kio, F. E. (2019). Colour vision deficiency among students in Lagos State, Nigeria. *African Health Sciences*, *19*(2), 2230. https://doi.org/10.4314/ahs.v19i2.48

Oliphant, D., & Hovis, J. K. (1998). Comparison of the D-15 and City University (second) color vision tests. *Vision Research*, *38*(21), 3461–3465. https://doi.org/10.1016/S0042-6989(98)00117-5

Olshausen, B. A., & Field, D. J. (2000). Vision and the coding of natural images.

American Scientist, *88*(3), 238–245. https://doi.org/10.1511/2000.3.238

Ou, L., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application*, *29*(3), 232–240. https://doi.org/10.1002/col.20010

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences*, *107*(19), 8877–8882. https://doi.org/10.1073/pnas.0906172107

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*, 77–107. https://doi.org/10.1146/annurev-psych-120710-100504

Panorgias, A., Kulikowski, J. J., Parry, N. R. A., McKeefry, D. J., & Murray, I. J. (2012). Phases of daylight and the stability of color perception in the near peripheral human retina. *Journal of Vision*, *12*(3), 1. https://doi.org/10.1167/12.3.1

Parry, N. (2015). Color vision deficiencies. In A. J. Elliot, A. Franklin, & M. D. Fairchild (Eds.), *Handbook of Color Psychology* (pp. 216–242). Cambridge University Press. https://doi.org/10.1017/CBO9781107337930.011

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, *18*(1), 79–85.

Pazda, A. D., Elliot, A. J., & Greitemeyer, T. (2012). Sexy red: Perceived sexual

receptivity mediates the red-attraction relation in men viewing woman. *Journal of Experimental Social Psychology*, *48*(3), 787–790.

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G., & Hurlbert, A. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0087989

Pease, P. L., & Allen, J. (1988). A new test for screening color vision: Concurrent validity and utility. *American Journal of Optometry and Physiological Optics*, *65*(9), 729–738. https://doi.org/10.1097/00006324-198809000-00007

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Penacchio, O., & Wilkins, A. J. (2015). Visual discomfort and the spatial distribution of Fourier energy. *Vision Research*, *108*, 1–7. https://doi.org/10.1016/j.visres.2014.12.013

Peterzell, D. H., Werner, J. S., & Kaplan, P. S. (1995). Individual differences in contrast sensitivity functions: Longitudinal study of 4-, 6- and 8-month-old human infants. *Vision Research*, *35*(7), 961–979. https://doi.org/10.1016/0042-6989(94)00117-5

Philipp, B. (2009). CircStat: a MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*. https://doi.org/10.18637/jss.v031.i10

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, *6*(39), eabb5053. https://doi.org/10.1126/sciadv.abb5053

Quinn, P. C., & Bhatt, R. S. (1998). Visual pop-out in young infants: Convergent evidence and an extension. *Infant Behavior and Development*, *21*(2), 273–288.

Quinn, P. C., Kelly, D. J., Lee, K., Pascalis, O., & Slater, A. M. (2008). Preference for attractive faces in human infants extends beyond conspecifics. *Developmental Science*, *11*(1), 76–83. https://doi.org/10.1111/j.1467-7687.2007.00647.x

Rabin, J., Gooch, J., & Ivan, D. (2011). Rapid quantification of color vision: The cone contrast test. *Investigative Ophthalmology and Visual Science*, *52*(2), 816–820. https://doi.org/10.1167/iovs.10-6283

Ramachandran, N., Wilson, G. A., & Wilson, N. (2014). Is screening for congenital colour vision deficiency in school students worthwhile? A review a. *Clinical and Experimental Optometry*, *97*(6), 499–506. https://doi.org/10.1111/cxo.12187

Ravens-Sieberer, U., Erhart, M., Rajmil, L., Herdman, M., Auquier, P., Bruil, J., Power, M., Duer, W., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J., & Group, the E. K. (2010). Reliability, construct and criterion validity of the KIDSCREEN-10 score: a short measure for children and adolescents' well-being and health-related quality of life. *Quality of Life Research*, *19*(10), 1487–1500.

https://doi.org/10.1007/s11136-010-9706-5

Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., Auquier, P., Power, M., Abel, T., & Czemy, L. (2005). KIDSCREEN-52 quality-of-life measure for children and adolescents. *Expert Review of Pharmacoeconomics & Outcomes Research*, *5*(3), 353–364.

Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M., & Klasen, F. (2014). The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *23*(3), 791–803. https://doi.org/10.1007/s11136-013-0428-3

Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, *21*(1–2), 97–117. https://doi.org/10.1163/156856808782713780

Redies, C., Hasenstein, J., & Denzler, J. (2007). Fractal-like image statistics in visual art: Similarity to natural scenes. *Spatial Vision*, *21*(1–2), 137–148. https://doi.org/10.1163/156856808782713825

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, *105*(3), 579–595. https://doi.org/10.1037/a0032690

Regan, B. C., Reffin, J. P., & Mollon, J. D. (1994). Luminance noise and the

rapid determination of discrimination ellipses in colour deficiency. *Vision Research*, *34*(10), 1279–1299. https://doi.org/10.1016/0042-6989(94)90203-8

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology*, *27*(12), 1825-1828.e3. https://doi.org/10.1016/j.cub.2017.05.044

Richardson, P., Saunders, K., & McClelland, J. (2008). Colour Vision Testing Made Easy; How Low Can You Go. *Optometry in Practice*, *9*, 17–24.

Rider, A. T., Coutrot, A., Pellicano, E., Dakin, S. C., & Mareschal, I. (2018). Semantic content outweighs low-level saliency in determining children's and adults' fixation of movies. *Journal of Experimental Child Psychology*, *166*, 293–309. https://doi.org/10.1016/j.jecp.2017.09.002

Rieth, C., & Sireteanu, R. (1994). Texture segmentation and "pop-out" in infants and children: The effect of test field size. *Spatial Vision*, *8*(2), 173–191. https://doi.org/10.1163/156856894X00323

Rinaldi, L. J., Smees, R., Alvarez, J., & Simner, J. (2020). Do the Colors of Educational Number Tools Improve Children's Mathematics and Numerosity? *Child Development*, *91*(4), e799–e813. https://doi.org/10.1111/cdev.13314

Robbins, S. L., Christian, W. K., Hertle, R. W., & Granet, D. B. (2003). Vision testing in the pediatric population. *Ophthalmology Clinics of North America*,

16(2), 253–267. https://doi.org/10.1016/s0896-1549(03)00007-5

Robles, K. E., Liaw, N. A., Taylor, R. P., Baldwin, D. A., & Sereno, M. E. (2020). A shared fractal aesthetic across development. *Humanities and Social Sciences Communications*, *7*(1), 158. https://doi.org/10.1057/s41599-020-00648-y

Rogers, M. R., Witzel, C., Rhodes, P., & Franklin, A. (2020). Color constancy and color term knowledge are positively related during early childhood. *Journal of Experimental Child Psychology*, *196*, 104825. https://doi.org/10.1016/j.jecp.2020.104825

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7*(2), 17. https://doi.org/10.1167/7.2.17

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Rovee-Collier, C., Hankins, E., & Bhatt, R. (1992). Textons, visual pop-out effects, and object recognition in infancy. *Journal of Experimental Psychology: General*, *121*(4), 435–445. https://doi.org/10.1037/0096-3445.121.4.435

Ruderman, D. L., Cronin, T. W., & Chiao, C.-C. (1998). Statistics of cone responses to natural images: implications for visual coding. *Journal of the Optical Society of America A*, *15*(8), 2036. https://doi.org/10.1364/josaa.15.002036

Saffran, J. R. (2020). Statistical Language Learning in Infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant Statistical Learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Salapatek, P. (1975). Pattern perception in early infancy. In L. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. 1, pp. 133–248). Academic Press.

Sanderson, C., & Curtin, R. (2016). Armadillo: a template-based C++ library for linear algebra. *The Journal of Open Source Software*, *1*(2), 26. https://doi.org/10.21105/joss.00026

Sanderson, C., & Curtin, R. (2018). A user-friendly hybrid sparse matrix class in C++. In J. H. Davenport, M. Kauers, G. Labahn, & J. Urban (Eds.), *Mathematical Software- ICMS 2018* (pp. 422–430). Springer. https://doi.org/10.1007/978-3-319-96418-8_50

Savazzi, F., Massaro, D., Di Dio, C., Gallese, V., Gilli, G., & Marchetti, A. (2014). Exploring responses to art in adolescence: A behavioral and eye-tracking study. *PLoS ONE*, *9*(7). https://doi.org/10.1371/journal.pone.0102888

Scherf, K., Behrmann, M., Kimchi, R., & Luna, B. (2009). Emergence of Global Shape Processing Continues Through Adolescence. *Child Development*, *80*(1), 162–177. https://doi.org/10.1111/j.1467-8624.2008.01252.x

Scherf, K., & Scott, L. (2012). Connecting developmental trajectories: Biases in face processing from infancy to adulthood. *Developmental Psychobiology*, *54*(6), 643–663.

Seshadri, J., Christensen, J., Lakshminarayanan, V., & Bassi, C. J. (2005). Evaluation of the new web-based “Colour Assessment and Diagnosis” test. *Optometry and Vision Science*, *82*(10), 882–885. https://doi.org/10.1097/01.opx.0000182211.48498.4e

Sharpe, L., Stockman, A., Jägle, H., & Nathans, J. (1999). Opsin genes, cone photopigments, color vision, and color blindness. In K. Gegenfurtner & L. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 3–51). Cambridge University Press.

Shepard, R. N. (1992). The perceptual organization of colors: an adaptation to regularities of the terrestrial world? In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 495–532). Oxford University Press.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology*, *59*, 143–166. https://doi.org/10.1146/annurev.psych.59.103006.093619

Shinomori, K., Panorgias, A., & Werner, J. S. (2016). Discrimination thresholds of normal and anomalous trichromats: Model of senescent changes in ocular media density on the Cambridge Colour Test. *Journal of the Optical Society of America A*, *33*(3), A65-76. https://doi.org/10.1364/josaa.33.000a65

Shrestha, R. K., & Shrestha, G. S. (2015). Assessment of color vision among school children: A comparative study between the Ishihara test and the Farnsworth D-15 test. *Journal of the Nepal Medical Association*, *53*(200), 266–269. https://doi.org/10.31729/jnma.2743

Shute, R., & Westall, C. (2000). Use of the Mollon-Reffin minimalist color vision test with young children. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, *4*(6), 366–372. https://doi.org/10.1067/mpa.2000.110335

Simion, F., & Giorgio, E. Di. (2015). Face perception and processing in early infancy: inborn predispositions and developmental changes. *Frontiers in Psychology*, *6*, 969. https://doi.org/10.3389/fpsyg.2015.00969

Simion, F., Leo, I., Turati, C., Valenza, E., & Dalla Barba, B. (2007). How face specialization emerges in the first months of life. *Progress in Brain Research*, *164*, 169–185. https://doi.org/10.1016/S0079-6123(07)64009-6

Simion, F., Macchi Cassia, V., Turati, C., & Valenza, E. (2001). The origins of face perception: specific versus non-specific mechanisms. *Infant and Child Development: An International Journal of Research and Practice*, *10*(1-2), 59–65.

Simoncelli, E. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, *13*(2), 144–149. https://doi.org/10.1016/S0959-4388(03)00047-3

Simoncelli, E., & Olshausen, B. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Skelton, A., Catchpole, G., Abbott, J. T., Bosten, J. M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(21), 5545–5550. https://doi.org/10.1073/pnas.1612881114

Skelton, A., & Franklin, A. (2020). Infants look longer at colours that adults like when colours are highly saturated. *Psychonomic Bulletin and Review*, *27*(1), 78–85. https://doi.org/10.3758/s13423-019-01688-5

Skelton, A., Franklin, A., & Bosten, J. (2021). Color vision is aligned with natural scene statistics at 4 months of age. *Manuscript in Preparation*.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2008). A Motivational Perspective on Engagement and Disaffection: Conceptualization and Assessment of Children's Behavioral and Emotional Participation in

Academic Activities in the Classroom. *Educational and Psychological Measurement*, *69*(3), 493–525. https://doi.org/10.1177/0013164408323233

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Smith, L., Yu, C., & Pereira, A. (2011). Not your mother's view: the dynamics of toddler visual experience. *Developmental Science*, *14*(1), 9–17. https://doi.org/10.1111/j.1467-7687.2009.00947.x

Smith, L., Yu, C., Yoshida, H., & Fausey, C. (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *Journal of Cognition and Development : Official Journal of the Cognitive Development Society*, *16*(3), 407–419. https://doi.org/10.1080/15248372.2014.933430

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

Smith, V., & Pokorny, J. (1996). The design and use of a cone chromaticity space: A tutorial. *Color Research & Application*, *21*(5), 375–383. https://doi.org/10.1002/(SICI)1520-6378(199610)21:5<375::AID-COL6>3.0.CO;2-V

Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society*

of London. Series B, Biological Sciences, *360*(1458), 1329–1346. https://doi.org/10.1098/rstb.2005.1633

Spears, W. C. (1964). Assessment of visual preference and discrimination in the four-month-old infant. *Journal of Comparative and Physiological Psychology*, *57*(3), 381.

Spehar, B., Clifford, C. W. G., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers and Graphics (Pergamon)*, *27*(5), 813–820. https://doi.org/10.1016/S0097-8493(03)00154-7

Spehar, B., & Taylor, R. P. (2013). Fractals in art and nature: why do we like them? In B. E. Rogowitz, T. N. Pappas, & H. de Ridder (Eds.), *Human Vision and Electronic Imaging XVIII* (Vol. 8651, p. 865118). https://doi.org/10.1117/12.2012076

Spehar, B., Walker, N., & Taylor, R. P. (2016). Taxonomy of Individual Variations in Aesthetic Responses to Fractal Patterns. *Frontiers in Human Neuroscience*, *10*, 350. https://doi.org/10.3389/fnhum.2016.00350

Spehar, B., Wong, S., van de Klundert, S., Lui, J., Clifford, C. W. G., & Taylor, R. P. (2015). Beauty and the beholder: the role of visual sensitivity in visual preference. *Frontiers in Human Neuroscience*, *9*, 514. https://doi.org/10.3389/fnhum.2015.00514

Stephens, B. R., & Banks, M. S. (1987). Contrast discrimination in human infants. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(4), 558–565. https://doi.org/10.1037/0096-1523.13.4.558

Steward, J., & Cole, B. (1989). What do color vision defectives say about everyday tasks? *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *66*(5), 288–295. https://doi.org/10.1097/00006324-198905000-00006

Stewart-Brown, S. L., & Haslum, M. (1988). Screening of vision in school: Could we do better by doing less? *British Medical Journal*, *297*(6656), 1111–1113. https://doi.org/10.1136/bmj.297.6656.1111

Stockman, A., MacLeod, D. I. A., & Vivien, J. A. (1993). Isolation of the middle- and long-wavelength-sensitive cones in normal trichromats. *Journal of the Optical Society of America A*, *10*(12), 2471–2490. https://doi.org/10.1364/JOSAA.10.002471

Stoianov, M., de Oliveira, M. S., dos Santos Ribeiro Silva, M. C. L., Ferreira, M. H., de Oliveira Marques, I., & Gualtieri, M. (2019). The impacts of abnormal color vision on people's life: an integrative review. *Quality of Life Research*, *28*(4), 855–862. https://doi.org/10.1007/s11136-018-2030-1

Studente, S., Seppala, N., & Sadowska, N. (2016). Facilitating creative thinking in the classroom: Investigating the effects of plants and the colour green on visual and verbal creativity. *Thinking Skills and Creativity*, *19*, 1–8. https://doi.org/10.1016/j.tsc.2015.09.001

Suero, M. I., Pérez, Á. L., Díaz, F., Montanero, M., Pardo, P. J., Gil, J., & Palomino, M. I. (2005). Does Daltonism influence young children's learning? *Learning and Individual Differences*, *15*(2), 89–98. https://doi.org/10.1016/j.lindif.2004.08.002

Sugita, Y. (2009). Innate face processing. *Current Opinion in Neurobiology*, *19*(1), 39–44. https://doi.org/https://doi.org/10.1016/j.conb.2009.03.001

Sullivan, K. (2011). Colour-blind children. *Early Years Educator*, *12*(11), 21–23. https://doi.org/10.12968/eyed.2011.12.11.21

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), 2879–2891. https://doi.org/10.1016/S0042-6989(02)00333-4

Tagarelli, A., Piro, A., Tagarelli, G., Lantieri, P. B., Risso, D., & Olivieri, R. L. (2004). Colour blindness in everyday life and car driving. *Acta Ophthalmologica Scandinavica*, *82*(4), 436–442. https://doi.org/10.1111/j.1395-3907.2004.00283.x

Tanabe, S., Ichikawa, H., Hukami, K., & Kawakami, G. (1978). New pseudoisochromatic plates for congenital color vision defects. *Japanese Journal of Clinical Ophthalmology*, *32*(3), 479–487.

Tang, T., Álvaro, L., Alvarez, J., Maule, J., Skelton, A., Franklin, A., & Bosten, J. (2021). ColourSpot, a novel gamified tablet-based test for accurate diagnosis of color vision deficiency in young children. *Behavior Research Methods*, 1–13. https://doi.org/10.3758/s13428-021-01622-5

Taylor, C., Schloss, K., Palmer, S. E., & Franklin, A. (2013). Color preferences in infants and adults are different. *Psychonomic Bulletin & Review*, *20*(5), 916–922. https://doi.org/10.3758/s13423-013-0411-6

Taylor, W. (1970). Diagnosis of Colour-Vision Defects in Very Young Children.

The Lancet, *295*(7650), 781–782. https://doi.org/10.1016/s0140-6736(70)91014-7

Tekavčič Pompe, M. (2020). Color vision testing in children. *Color Research and Application*, *45*(5), 775–781. https://doi.org/10.1002/col.22513

Tekavčič Pompe, M., & Stirn Kranjc, B. (2012). Which psychophysical colour vision test to use for screening in 3–9 year olds? *Zdravstveni Vestnik*, *81*, 170–177.

Teller, D. (1998). Spatial and temporal aspects of infant color vision. *Vision Research*, *38*(21), 3275–3282. https://doi.org/10.1016/S0042-6989(97)00468-9

Teller, D., & Bornstein, M. (1987). Infant color vision and color perception. *Handbook of Infant Perception*, *1*, 185–236.

Teller, D., Civan, A., & Bronson-Castain, K. (2004). Infants' spontaneous color preferences are not due to adult-like brightness variations. *Visual Neuroscience*, *21*(3), 397–401. https://doi.org/10.1017/S0952523804213360

Teller, D., McDonald, M., Preston, K., Sebris, S., & Dobson, V. (1986). Assessment of Visual Acuity in Infants and Children; the Acuity Card Procedure. *Developmental Medicine & Child Neurology*, *28*(6), 779–789. https://doi.org/10.1111/j.1469-8749.1986.tb03932.x

Teller, D., Pereverzeva, M., & Civan, A. (2003). Adult brightness vs. luminance as models of infant photometry: Variability, biasability, and spectral

characteristics for the two age groups favor the luminance model. *Journal of Vision*, *3*(5), 2. https://doi.org/10.1167/3.5.2

The KIDSCREEN Group Europe. (2006). *The KIDSCREEN Questionnaires - Quality of life questionnaires for children and adolescents*. Pabst Science Publishers.

Thomas, B. A. W. M., Kaur, S., Hairol, M. I., Ahmad, M., & Wee, L. H. (2018). Behavioural and emotional issues among primary school pupils with congenital colour vision deficiency in the Federal Territory of Kuala Lumpur, Malaysia: A case-control study. *F1000Research*, *7*, 1834. https://doi.org/10.12688/f1000research.17006.1

Thuline, H. C. (1964). Color-Vision Defects in American School Children. *JAMA*, *188*(6), 514–518. https://doi.org/10.1001/jama.1964.03060320036008

Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, *12*(2), 229–232. https://doi.org/10.1111/j.1475-1313.1992.tb00296.x

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, *14*(3), 391–412. https://doi.org/10.1088/0954-898X_14_3_302

Torrents, A., Bofill, F., & Cardona, G. (2011). Suitability of school textbooks for 5 to 7 year old children with colour vision deficiencies. *Learning and Individual Differences*, *21*(5), 607–612. https://doi.org/10.1016/j.lindif.2011.07.004

Tregillus, K. E. M., Isherwood, Z. J., Vanston, J. E., Engel, S. A., MacLeod, D. I. A., Kuriki, I., & Webster, M. A. (2021). Color Compensation in Anomalous Trichromats Assessed with fMRI. *Current Biology*, *31*(5), 936–942. https://doi.org/10.1016/j.cub.2020.11.039

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Treisman, A., Vieira, A., & Hayes, A. (1992). Automaticity and Preattentive Processing. *The American Journal of Psychology*, *105*(2), 341–362. https://doi.org/10.2307/1423032

Tummeltshammer, K., & Amso, D. (2018). Top-down contextual knowledge guides visual attention in infancy. *Developmental Science*, *21*(4), 1–9. https://doi.org/10.1111/desc.12599

Tummeltshammer, K., & Kirkham, N. (2013). Learning to look: probabilistic variation and noise guide infants' eye movements. *Developmental Science*, *16*(5), 760–771. https://doi.org/10.1111/desc.12064

Ugalahi, M., Fasina, O., & Ogun, O. A. (2016). Impact of congenital color vision defect on color-related tasks among Secondary School Students in Ibadan, Southwest Nigeria. *Nigerian Journal of Ophthalmology*, *24*(1), 20–24.

Ugalahi, M., Fasina, O., Ogun, O., & Ajayi, B. (2016). Prevalence of congenital colour vision deficiency among secondary school students in Ibadan, South-West Nigeria. *Nigerian Postgraduate Medical Journal*, *23*(2), 93–96. https://doi.org/10.4103/1117-1936.186301

Van Den Bergh, D., Van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E. J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E. J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *Annee Psychologique*, *120*(1), 73–96. https://doi.org/10.3917/anpsy1.201.0073

Van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E.-J. (2018). Bayesian inference for Kendall's rank correlation coefficient. *The American Statistician*, *72*(4), 303–308.

Van Hateren, J. H., & Ruderman, D. L. (1998). Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1412), 2315–2320. https://doi.org/10.1098/rspb.1998.0577

van Hof-van Duin, J., & Mohn, G. (1986). The development of visual acuity in normal fullterm and preterm infants. *Vision Research*, *26*(6), 909–916. https://doi.org/10.1016/0042-6989(86)90149-5

Velhagen, K. (1980). *Pflügerhaken-Tafeln zur Prüfung des Farbensinnes*. Thieme, Leipzig.

Ventura, D., Rodrigues, A., Moura, A., Vargas, A., Costa, M., de Souza, J., & Silveira, L. (2002). Color Discrimination Measured by the Cambridge Colour Vision Test (CCVT) in Children and Adults. *Investigative Ophthalmology & Visual Science*, *43*(13), 3796.

Verriest, G. (1982). Colour vision tests in children. *Documenta Ophthalmologica Proceedings Series*, *33*, 175–178.

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln* (Vol. 9). Voss.

Waggoner, T. (1994). *Colour vision testing made easy*. Waggoner Diagnostics.

Waggoner, T. (2017). *Frequently Asked Questions*. Colour Vision Testing. http://www.colorvisiontesting.com/color7.htm

Walker, F., Bucker, B., Anderson, N. C., Schreij, D., & Theeuwes, J. (2017). Looking at paintings in the Vincent Van Gogh Museum: Eye movement patterns of children and adults. *PLoS ONE*, *12*(6), 1–23. https://doi.org/10.1371/journal.pone.0178912

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001

Wass, S. V, & Smith, T. J. (2015). Visual motherese? Signal-to-noise ratios in toddler-directed television. *Developmental Science*, *18*(1), 24–37.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5), 3. https://doi.org/10.1167/11.5.3

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

Webster, M. A., & Mollon, J. D. (1997). Adaptation and the color statistics of natural images. *Vision Research*, *37*(23), 3283–3298. https://doi.org/10.1016/S0042-6989(97)00125-9

WHO Programme for the Prevention of Blindness and Deafness. (2003). *Consultation on development of standards for characterization of vision loss and visual functioning : Geneva, 4-5 September 2003*. World Health Organization. https://apps.who.int/iris/handle/10665/68601

Wilkinson, W. K. (1992). The cognitive and social-emotional correlates of color deficiency in children: a literature review and analysis. *Adolescence*, *27*(107), 603–611.

Witzel, C., Flack, Z., Sanchez-Walker, E., & Franklin, A. (2021). Colour category constancy and the development of colour naming. *Vision Research*, *187*, 41–54. https://doi.org/10.1016/j.visres.2021.05.008

Woldeamanuel, G. G., & Geta, T. G. (2018). Prevalence of color vision deficiency among school children in Wolkite, Southern Ethiopia. *BMC Research Notes*, *11*(1), 838. https://doi.org/10.1186/s13104-018-3943-z

World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Wright, W. D. (1952). The Characteristics of Tritanopia. *Journal of the Optical Society of America*, *42*(8), 509–521. https://doi.org/10.1364/JOSA.42.000509

Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae*. Wiley.

Yang, J., Kanazawa, S., & Yamaguchi, M. K. (2010). Perception of Munker-White illusion in 4-8-month-old infants. *Infant Behavior & Development*, *33*(4), 589–595. https://doi.org/10.1016/j.infbeh.2010.07.010

Yang, J., Kanazawa, S., Yamaguchi, M. K., & Motoyoshi, I. (2015). Pre-constancy Vision in Infants. *Current Biology*, *25*(24), 3209–3212. https://doi.org/10.1016/j.cub.2015.10.053

Young, T. (1802). II. 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. https://doi.org/10.1016/0042-6989(86)90143-4

Zeki, S., & Bartels, A. (1999). The clinical and functional measurement of cortical (in)activity in the visual brain, with special reference to the two subdivisions (V4 and V4 alpha) of the human colour centre. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *354*(1387), 1371–1382. https://doi.org/10.1098/rstb.1999.0485

Zemach, I., & Teller, D. (2007). Infant color vision: Infants' spontaneous color preferences are well behaved. *Vision Research*, *47*(10), 1362–1367. https://doi.org/10.1016/j.visres.2007.02.002

Żychaluk, K., & Foster, D. H. (2009). Model-free estimation of the psychometric

function. *Attention, Perception, and Psychophysics*, *71*(6), 1414–1425. https://doi.org/10.3758/APP.71.6.1414

Appendix

7.1 Supplementary Information for Paper 1

Table S1. A summary of existing adult color vision tests which have been used in children, outlining the test type, adult sensitivity and specificity values with their sample size and comparison test, rates of successful test completion in children with sensitivity and specificity if available, the recommended minimum age for test completion, and the limitations of each test.

Test name	Adult sensitivity/ specificity, numbers of adult test participants and comparison test	Details of tests on children including number of participants and sensitivity/specificity where available	Recommended minimum age for test	Limitations
Anomaloscope (Nagel, 1907)	1.00/1.00 Gold standard	7 year olds able to complete the test (Jurasevska et al., 2014)	7 years (Jurasevska et al., 2014; Verriest, 1982)	†, §, || Too demanding for young children
Pseudo-isochromatic				
Ishihara (Ishihara, 1917)	0.99 (Birch, 1997)/0.94 (Birch & McKeever, 1993) (CVD=401(Birch, 1997), CVN=471(Birch & McKeever, 1993)) Anomaloscope (Birch, 1997) Ishihara 1989 edition (Birch & McKeever, 1993; Ishihara, 1989)	79-97% of 2-6 year old children successfully completed the test (Choi & Hwang, 2009; Mäntyjärvi, 1991a)	7 years (Verriest, 1982)	†, ‡, § Requires number knowledge
Richmond Hardy-Rand-Rittler (Hardy et al., 1954)	1.00/0.98 (B. Cole et al., 2006) (CVD=100, CVN=50) Anomaloscope	N/A	3 years (Hardy et al., 1954)	†, ‡, § Requires shape knowledge
Standard Pseudo-isochromatic Plates (Ichikawa et al., 1979; Tanabe et al., 1978)	0.67/1.00 (Mäntyjärvi, 1987) (CVD=21, CVN=205) Anomaloscope	6-9 years (CVD=23, CVN=292) 0.87/0.99 (Haskett & Hovis, 1987) Ishihara 1974 edition (Ishihara, 1974)	N/A	†, § Requires number knowledge

Velhagen Pflügertrident Test (Velhagen, 1980)	0.61/1.00 (Mäntyjärvi, 1991b) (CVD=31, CVN=394) Anomaloscope	71% of 2-6 year old children successfully completed the test (Mäntyjärvi, 1991a)	4 years (Birch & Platts, 1993)	†, ‡, § Requires identification of a trident "E" figure orientated in different directions
KAMS (Jurasevska et al., 2014)	N/A	7-19 years 1.00/1.00 (Jurasevska et al., 2014) (CVD=8, CVN= 265) Anomaloscope	7 years (Jurasevska et al., 2014)	†, § Requires shape, symbol and number recognition
Arrangement				
Farnsworth Munsell 100-Hue (Farnsworth, 1943)	1.00/0.83 (Seshadri et al., 2005) (CVD=30, CVN=30) Anomaloscope	N/A	5 years (Kinnear & Sahraie, 2002)	†, ‡, || Performance is correlated with IQ (Cranwell et al., 2015)
Farnsworth Munsell D-15 (Linksz, 1966)	0.59/0.98 (Oliphant & Hovis, 1998) (CVD=70, CVN=81) Anomaloscope	8-15 years 0.92/0.99 (Shrestha & Shrestha, 2015) (CVD=24, CVN=1090) Ishihara	5 years (Kinnear & Sahraie, 2002)	† Performance is correlated with IQ (Cranwell et al., 2015)
Lanthony Desaturated D-15 (Lanthony, 1978)	0.79/1.00 (Marechal et al., 2018) (CVD=29, CVN=23) Anomaloscope	N/A	5 years (Lanthony, 1978)	†, ‡
Oddity				
City University Test (Fletcher, 1984)	0.95/0.83 (Oliphant & Hovis, 1998) (CVD=70, CVN=81) Anomaloscope	N/A	4 years (Fanlo Zarazaga et al., 2019)	†, ‡, § Requires knowledge of identifying orientation (up, down, left, right) verbally
Computerized				
Cambridge Colour Test (Regan et al., 1994)	0.94/0.92 (Shinomori et al., 2016) (CVD=32, CVN=162) Anomaloscope	Modified version tested in 2-7 year old children (Goulart et al., 2008)	7 years (Goulart et al., 2008; Ventura et al., 2002)	†, ‡, § Identifying the orientation of the Landolt C stimulus by pressing corresponding

				keys on a keyboard
Rabin Cone Contrast Test (Rabin et al., 2011)	1.00/1.00 (Rabin et al., 2011) (CVD=49, CVN=92) Anomaloscope	N/A	N/A	†, ‡, § Requires knowledge of letters
Colour Assessment and Diagnosis (CAD) (Barbur & Rodriguez-Carmona, 2015)	0.93/1.00 (Seshadri et al., 2005) (CVD=30, CVN=30) Anomaloscope	N/A	4 years (Barbur & Rodriguez-Carmona, 2015)	†, ‡

Note. **Definitions. Anomaloscope**: The anomaloscope is an optical instrument where individuals are asked to match different mixtures of red and green monochromatic light to different luminance levels of a yellow monochromatic light. It is the gold standard for assessing color vision; **CVD:** Participants with color vision deficiency (any CVD type, e.g. anomalous trichromacy, dichromacy); **CVN:** Participants with normal color vision; **N/A:** Not available; **Pseudo-isochromatic tests**: These tests have an array of colored dots that form a figure (digits, pathways, letters, animals or shapes) against an isoluminant background which individuals are asked to identify; **Arrangement tests**: Individuals are required to sort a set of colored stimuli by hue or saturation; **Oddity tests**: An odd-one-out task where individuals are asked to identify a colored target amongst other distractors; **Sensitivity**: The rate at which a diagnostic test identifies true positives (i.e. individuals with a condition are correctly identified). For example, in the table, against the comparison test (in this example, the anomaloscope), the Ishihara has a sensitivity of 0.99, indicating that 99% of individuals are correctly diagnosed as having a CVD (of any type) and 1% are false negatives (i.e., the Ishihara diagnosed the individual as having normal color vision (CVN) but the anomaloscope diagnosed the same individual as CVD); **Specificity**: The rate at which a diagnostic test to identifies true negatives (i.e. correctly identifies individuals the absence of a condition). For example, in the above table, when compared with the comparison test (in this example, the Ishihara 1989 edition), the Ishihara test has a specificity of 0.94, indicating that 94% of individuals of CVN individuals were correctly categorized as CVN, and 6% of individuals were false positives (i.e., where the Ishihara Unlettered diagnosed the individual as CVD but the Ishihara diagnosed the same individual as CVN).

Symbols. † Inaccessibility. The test is inaccessible for the public and/or requires specialized equipment and/or resources and/or a trained specialist administrator. **‡ Unknown validity**. The sensitivity and specificity values of the tests are unknown in children; **§ Unsuitability.** The test requires an understanding

of numbers, orientation, shapes, and/or animals or the task is so demanding that it is unsuitable for young children and children with additional educational needs; **|| Technical Limitations.** The test design or instructions do not address general and non-visual task demands that could make it difficult for young children to complete, for example, test duration is longer than 5 minutes, no gamification and no adaptive staircase procedure.

Distribution of errors on the Ishihara test for Unlettered Persons

Figure S1 shows the distribution of errors (≥1) on the Ishihara Unlettered test for participants in the discovery and validation cohorts. The distribution for participants that have at least one error looks bimodal, with a peak in errors at 4, but a second peak at 1. Participants with 1 or 2 errors were assigned to an 'Inconclusive' group and re-tested at a later date whenever practicable (*n*=6).

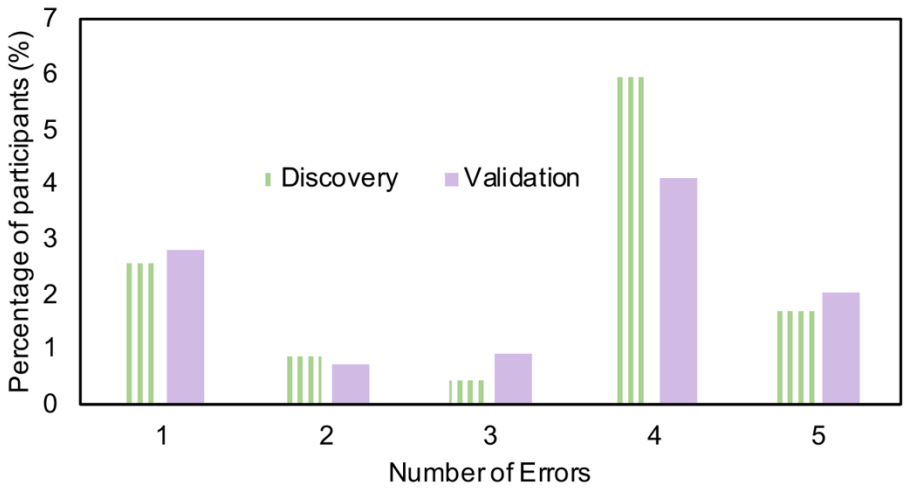


Figure S1. Percentages of participants who made errors on the Ishihara Unlettered test in the discovery cohort (N=236) and the validation cohort (N=536).

Of the 6 children in the discovery cohort re-tested, 4 children who made 1 or 2 errors on their first attempt later passed without error and were re-assigned to the control group and 2 children who again made 1 or 2 errors remained in the Inconclusive group.

Calibration

We investigated whether the protan, deutan and tritan thresholds measured from our participants differed between different iPad models. Mauchly's test of sphericity indicated the assumption of sphericity had been violated, ($\chi^2(2) = 0.54$, $p < 0.001$), and therefore a Greenhouse-Geisser correction was used. A mixed ANOVA (within-subjects factors: color confusion axis (protan, deutan, tritan)); between-subjects factor: iPad models (iPad Air 2, 2014; iPad Pro 9.7", 2016; iPad (5th Generation), 2017) showed no significant interaction between the color confusion axes and iPad model ($F_{2.74,\ 371.16} = 0.18$, $p = 0.90$) on thresholds, suggesting that protan, deutan and tritan thresholds do not significantly differ between iPad models.

Optimisation of fit parameters for categorising CVD versus normal color vision in the discovery cohort.

We did a systematic search of the optimal bandwidth and percentage correct criterion for defining threshold, which maximized the distance in threshold ratio between the CVD and control groups for the discovery cohort (Figure 5). The optimal parameters were determined as having a fixed bandwidth of 0.70 for the Model-free (Żychaluk & Foster, 2009) local linear fit, and a performance level of 0.21 (21%) on which to base threshold estimates. Figure S2 shows the separation between groups in threshold ratio (Z) against the two parameters (X and Y).

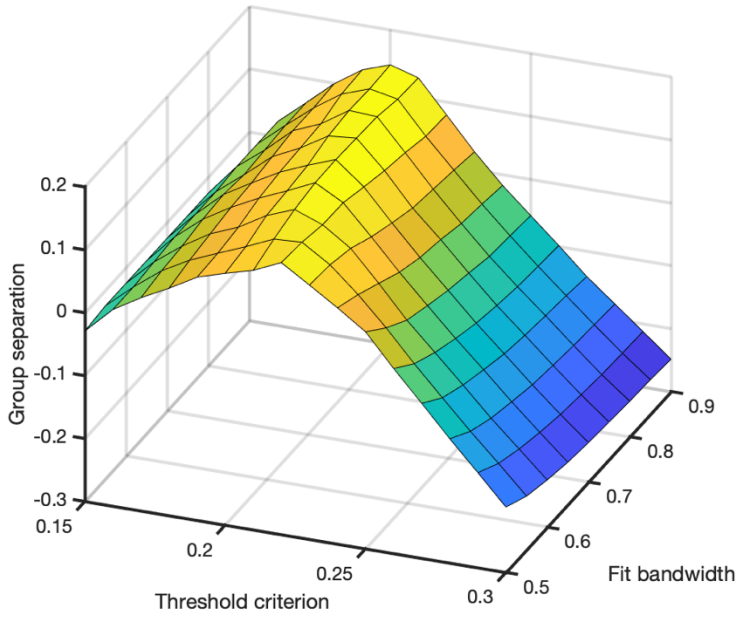


Figure S2. Results of a search for fit parameters that were optimal for distinguishing CVD from control participants. The Z-axis shows 'separation' in threshold ratio between the participant in the CVD group with the largest threshold ratio and the participant in the control group with the smallest threshold ratio. The two fit parameters that were varied were the proportion correct at which threshold was defined ('Threshold criterion) and the bandwidth ('Fit bandwidth') for the local linear fit made using Model-free (Żychaluk & Foster, 2009). The optimal parameters were defined at peak separation.

Raw protan and deutan thresholds

ColourSpot's tritan thresholds allowed us to quantify performance using the minimum ratio of tritan:protan thresholds or tritan:deutan thresholds. This offers an improvement in classification accuracy on many existing tests because non-visual influences on task performance (such as attention and task engagement) are factored out as they influence all three thresholds equally. The threshold ratio measure (Fig. 6) successfully separated those who were classified as CVD by the Ishihara Unlettered test and those

who were classified on that test as having normal color vision into two different groups. By contrast, if raw thresholds were used, as is the case for many existing tests, classification accuracy was poorer (Figure S3), since non-visual influences confer large individual differences in protan and deutan thresholds independently of color vision status.

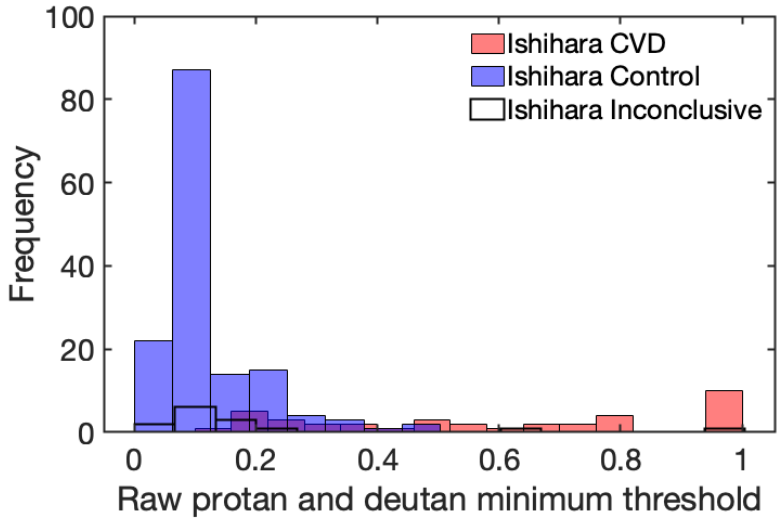


Figure S3. Histogram showing raw thresholds for participants in the discovery cohort grouped by color vision status assigned by the Ishihara Unlettered test. Plotted for each participant is the minimum of the protan or deutan threshold.

ColourSpot **Sensitivity and Specificity**

Discovery Cohort

Using the Ishihara Unlettered test as a pseudo gold standard, we identified a threshold ratio of 0.59 to be the classification criterion that distinguishes best between CVD and normal color vision, where a ratio smaller than 0.59 is indicative of a CVD. This threshold ratio was calculated by taking the average of the largest threshold ratio from the CVD group (0.52) and the smallest threshold ratio from the control group (0.66).

Excluding inconclusive participants, the criterion threshold ratio of 0.59 was applied to the discovery cohort (N=93) to estimate sensitivity and specificity values.

Table S2. Contingency table comparing the *ColourSpot* classification of CVD and normal color vision with that of the Ishihara Unlettered in the discovery cohort

ColourSpot	**Ishihara test for Unlettered Persons**	
	CVD	**Normal**
CVD	19	0
Normal	0	74
Total	**19**	**74**

Sensitivity=19/19=1.00

Specificity= 74/74= 1.00

Validation Cohort

For participants classified by the Ishihara Unlettered either as CVD or as having normal color vision (N=155), the threshold ratio value of 0.59 was independently applied to the validation cohort to estimate sensitivity and specificity for *ColourSpot*.

Table S3. Contingency table comparing the *ColourSpot* classification of CVD and normal color vision with that of the Ishihara Unlettered in the validation cohort.

ColourSpot	**Ishihara test for Unlettered Persons**	
	CVD	**Normal**
CVD	37	3
Normal	0	115
Total	**37**	**118**

Sensitivity= 37/37= 1.00

Specificity= 115/118= 0.97

For the validation cohort, *ColourSpot* classified 3 participants as having CVD which the Ishihara Unlettered had classified as having normal color vision. Two of these participants lie near the boundary between the normal and CVD group and cannot be classified with confidence. However, it appears that one of these participants (see Figure S4 for their psychometric function) was likely a false negative diagnosis by the Ishihara Unlettered (classified as having normal color vision but in fact has CVD). Based on their performance on *ColourSpot*, we can see that this participant had much poorer performance for protan and deutan targets than tritan, suggesting that a CVD diagnosis is appropriate.

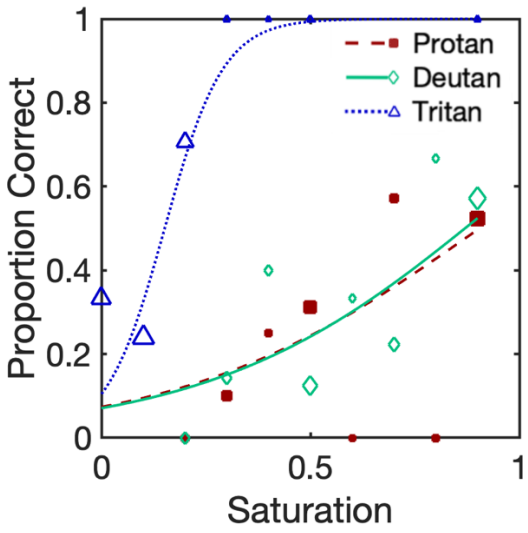


Figure S4. The psychometric functions of a participant in the validation cohort who was diagnosed as having normal color vision (made no errors) by the Ishihara Unlettered test but was diagnosed as CVD by *ColourSpot*.

Neitz Test of Color Vision

Only 111 boys (47.0%) completed the Neitz Test without any errors in the discovery cohort and 246 (45.9%) in the validation cohort. The breakdown of errors by type is given in Figure S5. The large number of errors is well above those expected from an 8% male prevalence for CVD, suggesting that the Neitz Test does not have good specificity for detecting CVD.

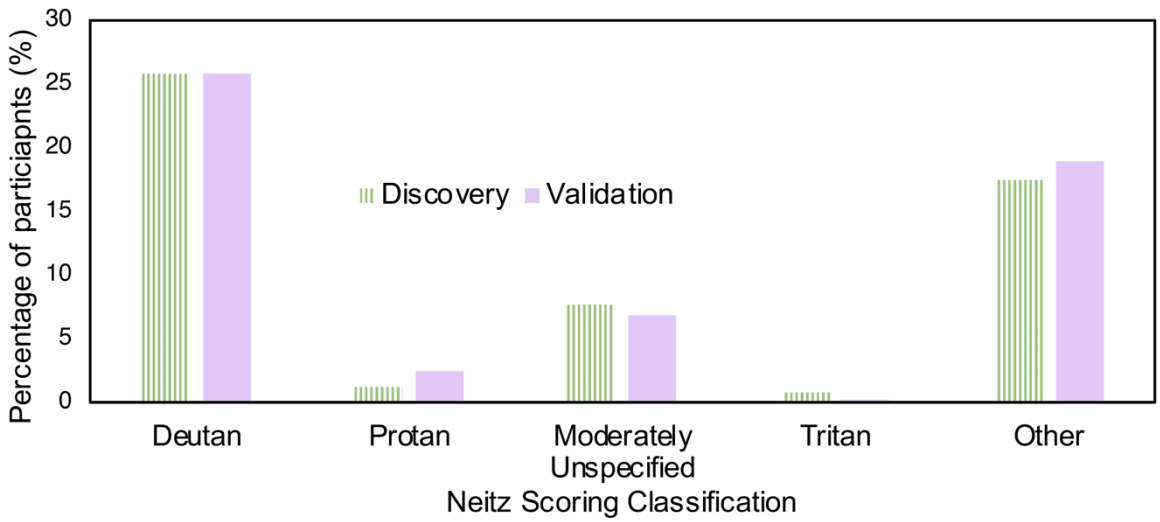


Figure S5. Percentage of participants that made errors of each type on the Neitz Test in the discovery and validation cohorts. 'Moderately Unspecified' is where CVD errors were made but the classification of error type cannot be made, and 'Other' is where errors were made that were not classified as a CVD error (e.g., an error due to misidentification of shape).

To further investigate the Neitz Test, the scoring classifications for the Neitz Test were compared to those of the Ishihara Unlettered test (Fig. S6). It can be seen that over 20% of children diagnosed as having normal color vision by the Ishihara Unlettered test (discovery and validation control groups) were classified as having deutan deficiencies by the Neitz Test. Additionally, about 17% of children who were diagnosed as having normal color vision by Ishihara Unlettered test made other errors on the Neitz Test. Furthermore, approximately 7% of children classified as CVD by the Ishihara Unlettered test were classified as having CVD that was "moderately unspecified" by the Neitz Test. These comparisons further suggest the Neitz Test overdiagnoses young

children with CVD.

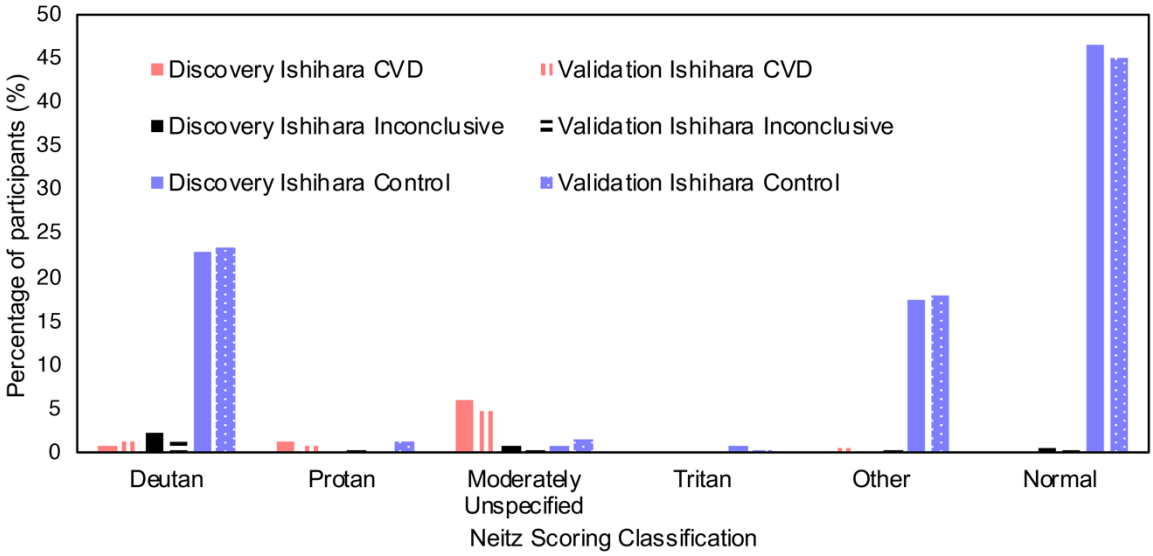


Figure S6. Scoring classifications on the Neitz Test against classifications made from the Ishihara Unlettered test in the discovery and validation cohorts.

7.2 Supplementary Information for Paper 3

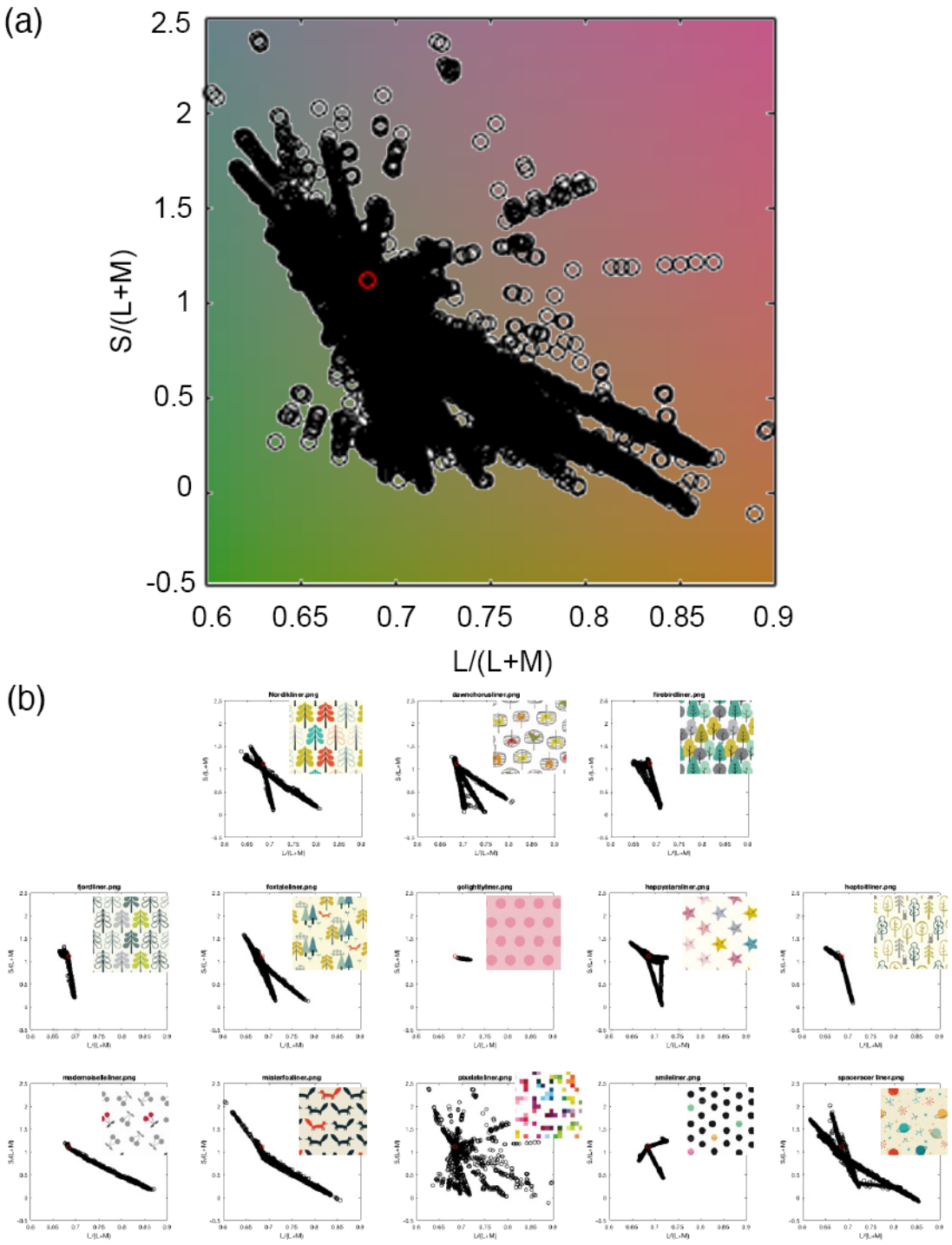
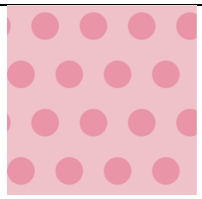


Figure S7. The colour distribution in (a) all the pattern images and (b) each pattern image plotted individually in cartesian space in the MacLeod-Boynton chromaticity diagram with the red circle as the white point (0.6852,1.1154).

Table S4. Categorising Cosatto Ltd patterns by the most dominant hue.

Hue types	**Most dominant hue patterns**
Pink-reddish	
Chartreuse-ish	
Cyan-blue-ish	
Achromatic