2021

Color Vision Testing with Readily Available Materials

Angel Perez
University of Alabama at Birmingham

COLOR VISION TESTING WITH READILY MADE MATERIALS

by

ANGELEA PEREZ

MICHAEL S. LOOP, COMMITTEE CHAIR
PATTI FUHR
TIM GAWNE
LEI LIU
LAWRENCE SINCICH

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2021

COLOR VISION TESTING WITH READILY AVAILABLE MATERIALS

ANGELEA PEREZ

VISION SCIENCE

ABSTRACT

The perception of color is a process by which the brain discriminates different light wavelengths stimulating the cone photoreceptors in the retina of the eye. In humans each cone contains a photopigment making it most sensitive to either short (red), medium (green), or long (blue) wavelengths of light. However, when one of the cone pigments is defective or missing, color discrimination is reduced, resulting in color vision deficiency due to a deficient sex-linked chromosome in most cases. Color vision deficiency has the potential to impede many everyday activities, interfere with the learning processes for children beginning at a very early age, and prohibit participation in numerous careers and occupations for which color recognition is critical. This study examined four economized tests we developed as possible screening tools for color vision deficiency (CVD) that can be used in educational environments where children could be assessed.

This study explored the ability to develop a valid and reliable color vision test using materials that are readily available and easily accessible and were compared to the widely used Hardy, Rand and Rittler Pseudoisochromatic Plates. Forty-nine subjects (35 normal and 14 color vision deficient) performed the HRR, crayon, color board, paint chips, and a psychophysical measure, the red test projected on a gray

background. The HRR had a 37% error rate with plate 7 for normals and 6% of normals and 10% of the CVDs made errors on the crayon test and does not discriminate well between normals and mild CVDs. Next, 17% (35 subjects) of the normal subjects made errors on the color board test and 57% (14 subjects) of the CVDs had 100% accuracy. With the paint chips test, normals identified all colors with 100% accuracy while only 36% (14 subjects) of the CVDs made errors. The normals performed well with the red test even with the smallest increment 0.75, while the CVDs performance was worse on average and scores for the smallest increment indicated that this population guessed for due to the inability to detect this increment.

Our findings indicate that both the color board and crayon tests were not reliable because of the desaturation of color in both. The paint color test results were slightly better; however, they issue with these tests are that they would need to be measured using a spectrometer to match the confusion line on the CIE Chromaticity diagram. The red test was determined to be consistent with the HRR and can detect most CVDs, and those with the most severe CVD in comparison to the normals do not appear to be the source for the difference in CVD variation levels.

Keywords: Color Vision Deficiency, Hardy-Rand Rittler, Protan, Deutan

DEDICATION

To my wonderful husband, Joe, who has been beside me through everything, thank you for your support, love, patience, and encouragement. I want to thank my son, Tristan for his understanding of my absence during this long road and giving me the motivation to reach for the stars. To my grandson, Lynken, who is a motivation within himself to be the best I can be.

ACKNOWLEDGEMENTS

I want to thank Dr. Michael Loop, my mentor and friend, who had enough faith in me to provide this research opportunity in the beginning of this long path. To Dr. Lawrence Sincich, who not only became a great mentor to me, but a very good friend. You believed in me and continuously encouraged me to press on to the end and guided me through the most critical time in the process.

I want to thank my committee members Dr. Gawne, Dr. Lui and Dr. Fuhr for your patience, efforts, and contributions. To Dr. Stefanie Varghese for your support and encouragement and always helping me out of a jam. Thank you to Dr. Ellen Bowman and Diane Ackerman, you never gave up on me and was always there when I needed a friend. And to Amanda Perez who never stop believing. To SLN for your support, you had no doubt about my ability to accomplish anything I put my mind to.

The National Leadership Consortium for Sensory Disabilities Organization (NLCSD) provided me with the greatest opportunity of all, the opportunity to be a doctoral student and provided me with amazing educational opportunities along the way. To the people of New Jersey who agreed to be my subjects, and to my fellows,who became my beloved friends. I want to thank The College of New Jersey and Dr. Susanne McCotter for allowing me to use a room in the college to test subjects.

A big thank you to Dr. Regina Bussing for your encouragement at the end stages and my future opportunities, and to all my friends and family who cheered me along this road.

Finally, to the University of Alabama at Birmingham School of Optometry, thank you for your support both in optometry school and as a graduate student. I will forever be a Blazer at heart.

"But they that wait upon the LORD shall renew their strength; they shall mount up with wingsas eagles, they shall run and not be weary, and they shall walk and not faint."
Isaiah 40:31

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

cd/m^2	Candelas per square meter
CIE	Commission Internationale de l'éclairage
CVD	Color Vision Deficiency
DPL	Digital Light Processing
FM-100	Farnsworth-Munsell 100 Hue Test
HRR	Hardy-Rand Rittler
LCD	Liquid Crystal Display
LGN	Lateral Geniculate Nucelli
L.M.S	Long, Medium, Short
PD	Protan/Deutan
RGB	Red/Green/Blue
SD	Standard Deviation
Vλ	Luminosity Function

CHAPTER 1

Introduction

Coren and Hakstian (1988) provided a list of specifications for developing a useful color vision test. This list is comparable to the standards I hope to achieve in developing a new test protocol that: (1) can be validated against standard laboratory and clinical tests for color vision deficiency, (2) must be statistically reliable, (3) is suitable for group administration, (4) is capable of measuring the broad range of sensory dysfunction, from normal through deficient, (5) can be applied across many age groups, (6) the test is brief and easy to comprehend, (7) is easily reproducible in format, requiring no special color plates or pictorial matter, (8) can produce results that will be meaningful in terms of the presence or absence of a color vision deficiency, and finally (9) is not dependent upon previous clinical diagnoses or direct knowledge of previous color-vison testing. In developing the tests described in this dissertation, we kept these specifications in mind, with the ultimate goal of creating assessments for primary use among school children from grades K-12. Before being used in that population, it was important to test them with adults who were known to be color vision deficient or normal, as an initial "proof of concept".

Humans possess an incredibly complex visual system allowing them to perceive colors and much else in the environment. The color vision of humans has been studied for well over 200 years (Young, 1802). Color vision is a crucial element of human vision, and performs an essential role in perception and communication in our world.

Color vision begins with the absorption of light in the retinal cone photoreceptors, where photopigments convert electromagnetic energy into electrical signals measurable as changes of cellular membrane potential. These signals are passed to bipolar cells and then transformed into action potentials via the ganglion cells in the retina, which sends information to the visual cortex by way of the lateral geniculate nucleus (LGN) in color-opponent channels characterized psychophysically, physiologically, and behaviorally (Gegenfurtner and Kiper, 2003). Chaparro et al. (1993) concluded that a colored stimulus is seen with a minimum of three times greater sensitivity than the best luminance stimulus. Sensitivity to color appears consistent with the high color contrast gain of midget ganglion cells of the retina and possibly compensates for any low chromatic contrasts found in nature (Watson, Barlow, and Robson, 1983). Humans have the ability to differentiate various colors and the foundation of color vision is fundamental to detecting color vision anomalies (Pasmanter and Munakomi, 2020). Those who have congenital color vision deficiency may experience problems with a number of occupations and adverse effects with everyday life. Testing color vision early in children is important in providing intervention and academic success.

The human retina contains three types of light sensitive cells: rod photoreceptors, cone photoreceptors, and the photosensitive retinal ganglion cells which mediate image-forming vision as well as non-image forming physiological responses to light (Schroeder, et al.) The rods and cones collect detailed information about spatial and temporal light levels in a scene and pass this on to the rest of the visual system which forms a representation of the external world.Vision in low lighting or at night is due to the rods while the cones contribute vision in room light and daylight (Neitz and Neitz,

2000). There are approximately 120 million rod photoreceptors, containing the photopigment rhodopsin, and they are responsible for what is known as scotopic vision. Their contribution to color vision is minimal. Cone photoreceptors number around six to seven million, are more centrally located in the retina, and are responsible for photopic vision and color sensitivity (Curcio et al, 1991). Normally, cone photoreceptors contain one of three different types of visual pigments with varying spectral sensitivity (**Figure 1.1**), the short wave or blue, the middle wave orgreen, and the long wave or red with greatest absorption at around 420 nm, 530 nm, and 560 nm, respectively (Yamaguchi, Motulsky, and Deeb, 1997).

Schwartz (2010) notes that the basic classification for abnormal color vision is divided into two categories, which are dichromacy and anomalous trichromacy. The characteristics of abnormal color vision are deficits in spectral sensitivity, color confusion lines, wavelength discrimination, and saturation sensitivity. We will outline each of these deficits in turn.

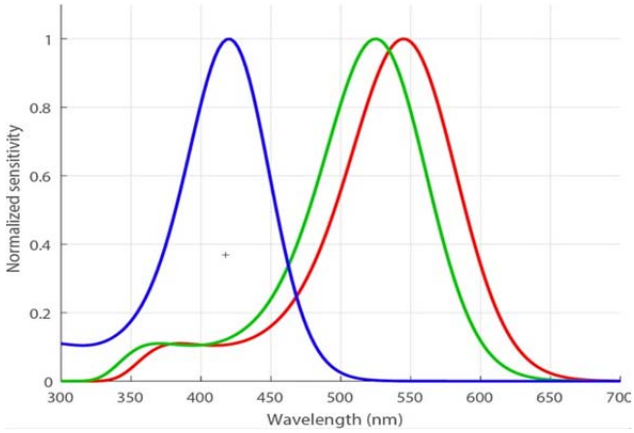


Figure 1.1. Normalized wavelength sensitivity of human cones. Data from Stockman and Sharpe (2000).

1.1 Types of Color Vision Deficiency

The most common color vision deficiency (CVD) is the congenital absence or alteration of a cone photopigment. It is a genetic defect and is much more common in men than women. Approximately 8% of the male population and 0.4% of female population are affected by this disorder. The most prevalent form of CVD is that of red-green deficiency which is caused by an absence of M-cones or L-cones (Schwartz and Krantz, 2015). Red-green color deficient dichromats (2% of all males) have the most extreme inherited red-green deficiency based on having just two cone types instead of the normal three cone types (red, green, blue). Dichromat color deficient people tend to confuse a large part of the spectrum and can match a specific portion, the neutral point, with white and any color with a mixture of two primary colors (Hecht and Shlaer, 1935).

Anomalous trichromats have a milder form of red-green color deficiency and are characterized by two types, protanomaly and deuteranomaly (Neitz and Neitz, 2000). This group is inclined to confuse narrower sections of the spectrum but require three primary colors to form a match to any color. People who are protanomalous are understood to have normal S and M pigment, but abnormal L pigment or in more specific terms, their 'L' pigment is shifted closer to their "M" pigment than in normals (Neitz and Neitz, 2000).

Deuteranomaly is not only the more common of the anomalous trichromacies, it is the most common color vision anomaly for those who have an inherited CVD, and is estimated to effect one out of twenty men. Even though it is based on three cone photopigments, deuteranomalous color defects are caused by various shifts in the wavelength of maximum sensitivity of the M photopigment (Neitz and Neitz, 2000).

Other less prominent color vision deficiencies are tritanopia which is a very rare form of color vision deficiency that is a non-sex linked (autosomal dominant) trait and is caused by short wavelength photopigment gene called OPN1SW located on chromosome 7. Blue Cone Monochromacy, another rare form of color vision deficiency that abolishes the function of both the long and medium wavelength photopigment genes and is X-linked recessive (Deeb, 2004).

The rarest form of color vision deficiency is referred to as achromatopia, often described as true colorblindness. Achromatopsia is either a partial or complete absence of color vision, indicating that only black, white, and grays can be perceived. Achromatopsia affects approximately 1 in 30,000 people throughout the world and is generally associated with other visual disorders. For the purpose of this study, we will concentrate solely on the most common forms of color vison deficiency, red-green color defects.

1.2 Genetic Causes of Color Deficiency

Most color vision deficiency is a sex-linked genetic trait. Females have two X-chromosomes, one from the mother and one from the father while the males have one X-chromosome and one Y-chromosome. The long and medium wavelength photopigment genes called OPN1MW and OPN1LW are located on the X-chromosome. The genetic locus control region regulates the activity of these two genes and only the two opsin pigment genes closest to the locus control region are active, allowing the photopigments to be expressed in the cones and so contributing to color vision (Deeb, 2004). Red-green color vision defects are inherited and because the mother of the affected child is a heterozygous, the chance of transmitting the altered gene in each birth is 50%.

Males who inherit this altered gene will be color deficient; females who inherit this altered gene will be carriers. Affected males transmit the altered gene to their daughters, who will be carriers (Deeb and Motulsky, 2005).

1.3 Elements of Color Vision Deficiency

In persons with trichromacy (normal color vision), the increment threshold of the color-opponent spectral sensitivity has three peaks with broad overlapping spectral absorption (Sperling and Harwerth, 1971; King-Smith and Carden, 1976). But, the

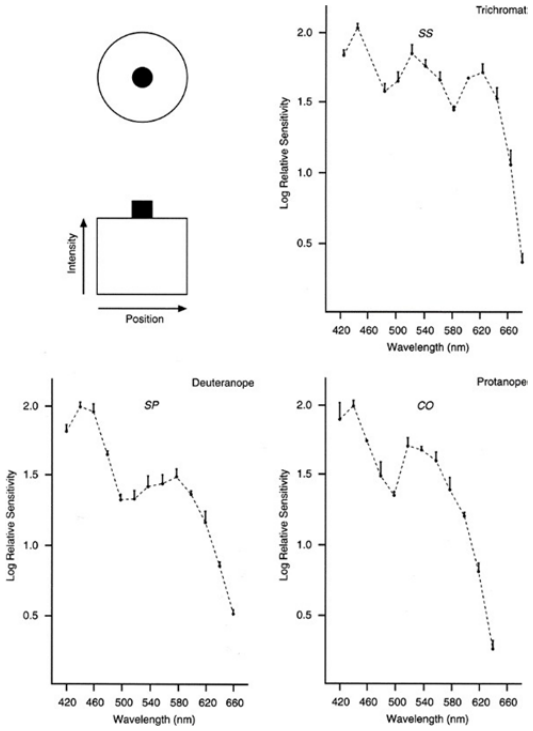
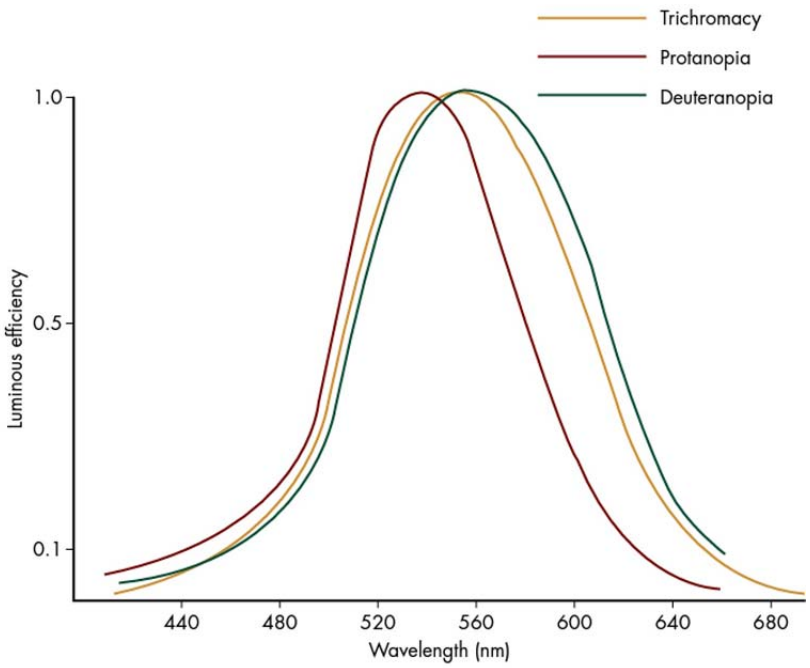


Figure1.2. Chromatic increment spectral sensitivity functions for a trichromat, deuteranope, and protanope. Stimulus conditions are given in the upper left. Adapted from Schwartz (1994).

results from people with dichromacy look very different (**Figure 1.2**). Those who are deuteranopes have two peaks parallel to the two photopigments, red and blue (long and medium cone photopigments). Protanopes also have two peaks, but they parallel with

green and blue photopigments (Verriest and Uvijls, 1977; Schwartz, 1994). When the red pigment is missing (protanopes), it causes the luminance efficiency curve to shift toward shorter wavelengths and when the green pigment is missing (deuteranopes) it causes the luminance efficiency curve to move toward longer wavelengths (**Figure 1.3**). For normal trichromacy, there is a broad peak around 555 nm, and in viewing the deuteranopia function peaks at a wavelength slightly longer than normal. However, protanopia is different in that the curved is narrower and shifted toward the shorter wavelengths which explains why these individuals most likely struggle with the ability to see red objects at low luminance (Hsia and Graham, 1957; Schwartz, 1994).



Wavelength discrimination of 490 nm for protanopes and deuteranopes seem to be well-developed, but wavelengths beyond 545 nm lead to an inability to discriminate between stimuli based on wavelength differences alone (Schwartz, 1994).

Figure1.3. Photopic luminance functions V_λ for trichromacy, deuteranopia, and protanopia. From Schwartz (2015) based on data from Hsia and Graham (1957).

Another measure of color deficiency appears with wavelength discrimination. This is typically measured by determining how large a spectral difference is needed to perceive that two wavelengths are not the same. For instance, when testing at 500 nm, a person with normal trichromacy would need a difference of about 2 nm to tell if the comparator is not 500 nm. **Figure 1.4** shows wavelength discrimination for protanopia, deuteranopia, and normal trichromacy. The wavelength discrimination for abnormal

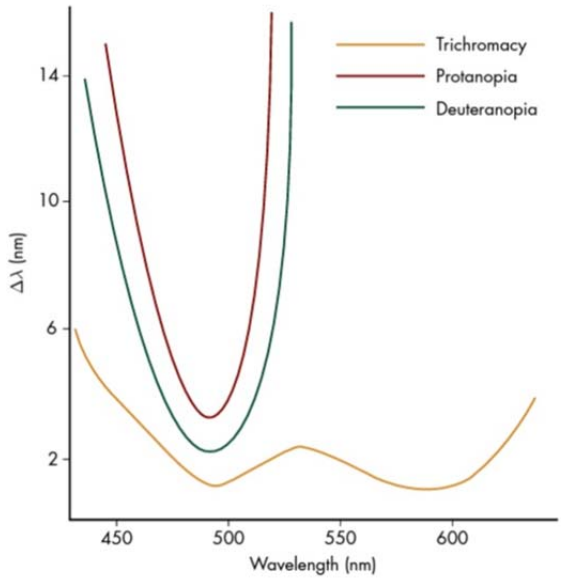


Figure 1.4. Wavelength discrimination functions for trichromacy, deuteranopia, and protanopia. From Schwartz (2015) based on data from Pitt (1935).

types of color vision approach the normal values only near 490 nm; with wavelengths longer than 545 nm, discrimination by dichromats is unable to occur (Pitt, 1935; Wright, 1952).

The sensation of whether a color is saturated is also a function of wavelength. Saturation of a color is measured by its purity and dominant wavelength. In normal trichromats, the point with least perceived color saturation is near 570 nm, which is

yellowish (Schwartz 1994). For dichromats, the points near 490-500 nm look gray or white, which indicates zero color saturation (**Figure 1.5A**). For anomalous trichromats, color deficits are less severe (**Figure 1.5B**). They do not have an absolute neutral point, but they do have saturation minimal at wavelengths that are near the dichromatic neutral points. In general, for both types of abnormal color vision, color percepts are less saturated, as shown by their saturation curves being below normals across most of the spectrum.

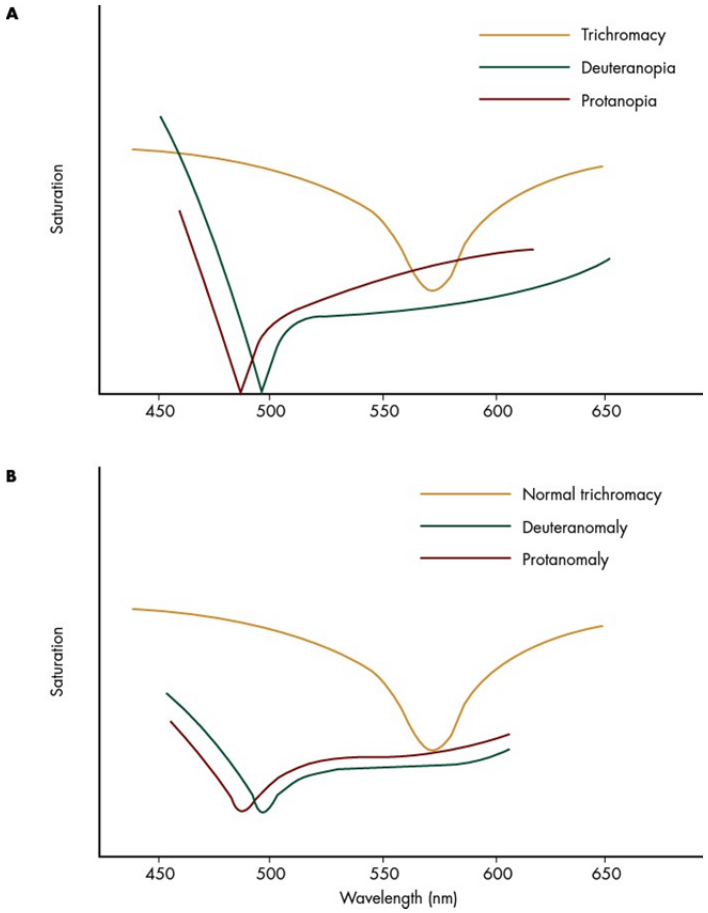


Figure 1.5. Wavelength saturation functions for dichromats (A) and anomalous trichromats (B). Yellow curves show normal trichromacy in both graphs. Note the intersections of the deuteranopic and protanopic functions with the abscissa at 498 and 492 nm, respectively. These wavelengths represent the dichromatic neutral points, which appear colorless (white or gray). Anomalous trichromats have minima in the same wavelength region, but the percepts are not colorless. From Schwartz (2015; Chapanis)

1.4 Neutral Point

The neutral point is a single wavelength of light that appears achromatic (gray, black or white) for subjects who are dichromats. This has been rendered in **Figure 1.6**, which illustrates that no intensity of spectral light could induce a color percept. For a protanope this neutral point is approximately 490 nm, and for a deuteranopes it is approximately 500 nm (Benson, 2013). When presented on the CIE chromaticity chart (**Figure 1.7**), there are lines that run through the graph that give an indication of the colors dichromats cannot discriminate, these are called confusion lines. An important confusionline passes through the white point on the graph (x=0.33, y=0.33) and touches the spectrum neutral point, while the other end of the line touches the non-spectral purple line.

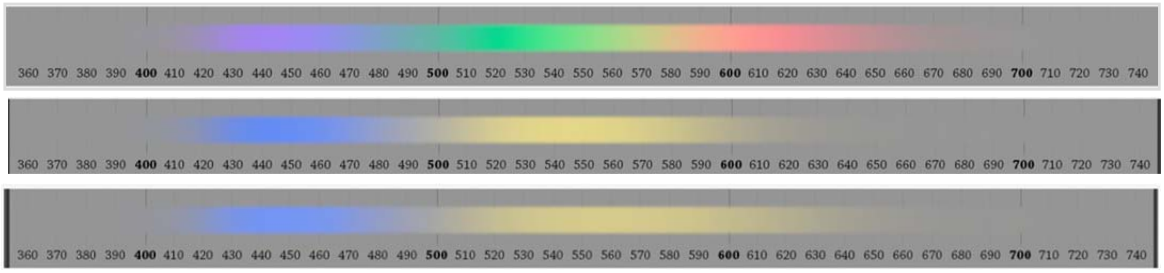


Figure 1.6. Rendered appearance of color saturation across the visible spectrum. From top to bottom, appearance of the equal energy spectrum for individuals with normal trichromacy (top), protanopia (middle), and deuteranopia (bottom). The spectrum for protanopia and deuteranopia divide into blue and yellow regions across their respective neutral points which appear gray. Color appearance created using the "Color-Blindness" proofing operation in Adobe Photoshop CS6.

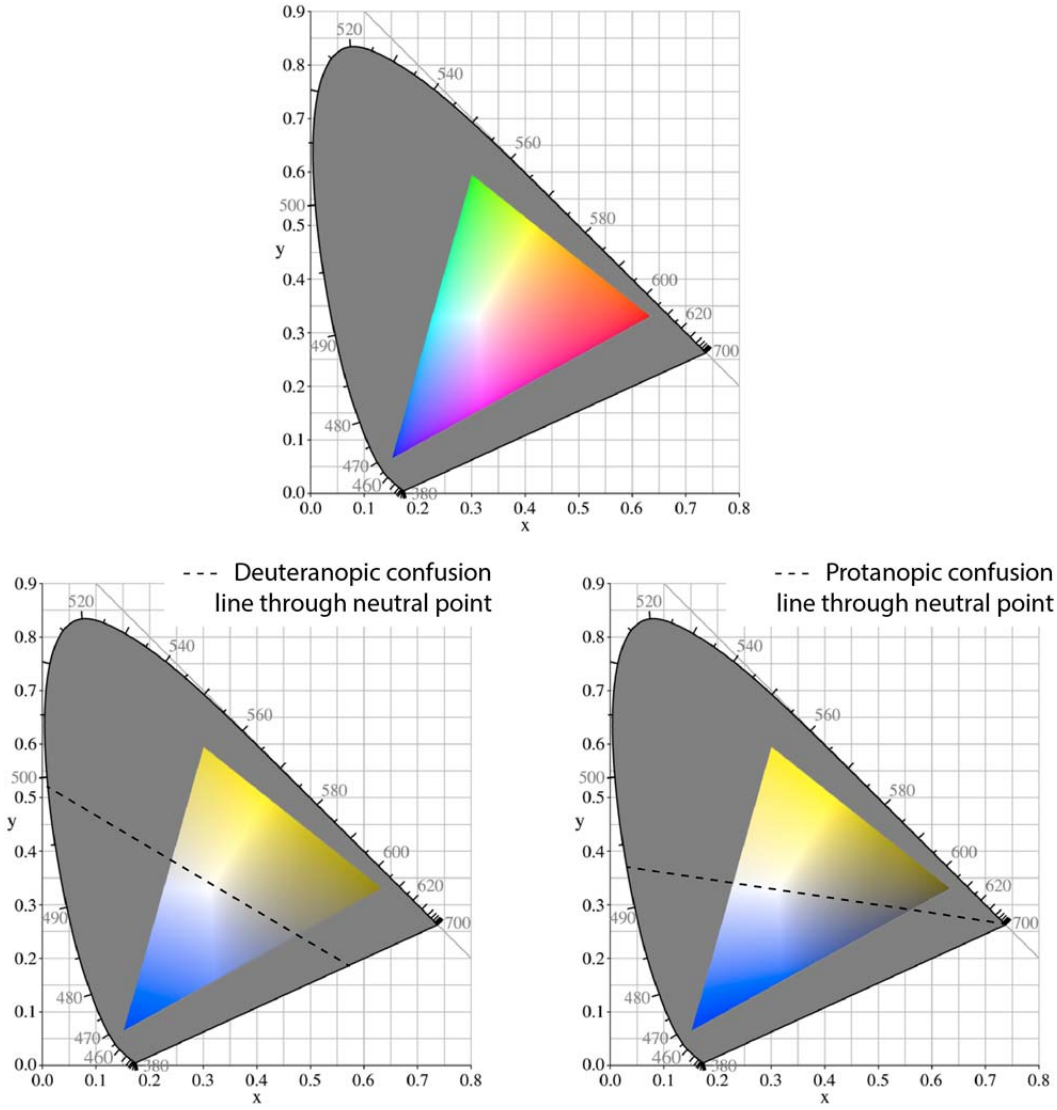


Figure 1.7. CIE chromaticity diagrams rendering the RGB gamut for normal trichromacy (top), deuteranopia (bottom left), and protanopia (bottom right). The deuteranopic and protanopic gamuts are divided into blue and yellow regions separated by a white neutral point, through which a confusion line passes. Color appearance created using the "Color-Blindness" proofing operation in Adobe Photoshop CS6

1.5 Color Vision Deficiency and Children

It has long been debated whether the inability to see color has adverse effects on those with CVD leading to driving risks, occupational hazards and limitations, and

educational performance, or whether they have developed methods for accommodating for their loss regardless of the severity. Cole (2004) interviewed 102 color vision deficient participants to assess the behavioral effects their deficiency had on their everyday activities. Of the 102 interviewees, only seven reported that they did not encounter any issues with daily activities.

Many people with CVD have complained of their inability to differentiate cooked and uncooked meat, raw or ripe vegetables and fruits, red and green electrical wiring, and red, green, and yellow traffic signals (Cole, 2004). Another study found that people with CVD take as much as 42% to 98% more time responding to color signals while driving (Whillans and Allen, 1992). In 2010, a group of scientists examined the effects of color vision deficiency among medical students from Nepal Medical College and Teaching Hospital. The students with CVD (57.0% were protanopic and 43.0% were deuteranopic) could not recognize changes to the body's color, body products such as blood, bile, vomit, and mouth or throat conditions, and color-indexed charts and test strips such as those used for blood and urine samples (Pamanik, Sherpa, and Shrestha, 2010). As for impacts to occupations, Blais (2010) listed the top 100 jobs where color vision is crucial in his Richmond Products CVD tutorial for Optometry and Ophthalmology, stating "depending on the nature of the business, an employee's ability to discern certain colors can be critical". He also goes on to state, "An important reminder to both employers and their employees – the American Disability Act of 1990 states that an individual must be able to perform the essential task with or without correction without significant risk or increased threat to the individual and the workplace." Color vision testing is essential in

order to evaluate a person's ability to perform certain tasks and occupations adequately and safely not only for themselves, but others as well.

Ling and Dain (2008) report that there are only a handful of studies concerning the perception of color in children in their early school years, as well as a shortage of adequate testing protocol suited for children. Another issue is the lack of any requirement in school age children to have their color vision screened in American and some European countries like Germany. The National Center for Children's Vision and Eye Health (2020) found that not all states require vision screenings for pre-school and school age children. Currently, 78% of the states require school-age children to be screened and 51% require pre-school children to be screened, and only 16 states require color vision screenings for these age groups (Suckow and Watson, 2002). It is also worth noting that among screening standards for driving, Massachusetts is one of the few states that requires color vision testing as one of their standards for receiving a driver's license (Steinkuller, 2010).

Not all studies agree about the effect of CVD on daily living. Ramachandran, Wilson, and Wilson, (2014) claimed in their review that having a color vision deficiency does not increase the risk for automobile accidents, that there is no correspondence between color vision deficiency and educational achievement, and they are skeptical about the value of screening students for future occupational purposes. Furthermore, they felt that even though people with CVD perform more poorly in specific occupational tasks where color is the primary source used to complete these tasks, they say that relevance in 'real world' operations are limited and that employees with color vision

deficiency most likely compensate in some fashion, or use the only recognized intervention, tinted lens, to help CVD employees manage their color perception. On the other hand, they consistently found that the choice of occupations was compromised by the knowledge of defective color vision. They also reported that those with CVD performed substantially poorer in job performance where color was the determining factor.

In a rebuttal to this opinion, Long, Honson, Katalinic, and Dain (2015), argued that waiting until a person reaches adulthood to test color vision is entirely too late because there are potential investments, both emotional and financial, that have already been committed to a particular career field. Ramachandran, Wilson, and Wilson (2014) concluded that there is a psychological impact that comes with a delayed diagnosis and that early counseling is underestimated, leaving people in their teens and early twenties in a sense of shock when given the diagnosis of color vision deficiency as they began planning for their future careers. They go on to add that other emotions surface such as grief, disbelief, and anger. Long, Honson, Katalinic, and Dain (2015) believe that the earlier a person can be tested for CVD, the better the chances of having classroom accommodations as well as career guidance.

There are students who do not know they have a color deficiency, and are at risk of being labeled as "learning disabled" by education professionals who are unaware of CVD. For instance, color as an identifier, particularly in school, is used systematically to group objects, ideas, and areas in the classroom. It has been used in the academic tasks over the course of decades to classify everything from English phonemes to establishing a Cuisenaire method of presenting mathematical relationships. Most

compelling was a research study produced by Gallo, Panza, and Viviani (1998) testing 82 students with color vision deficiency who were found to be unquestionably inferior in academic achievement to those children matched by age and class who had normal color vision. One student recalled an experience with his kindergarten teacher who drilled him on color names and was left feeling anxious and ashamed from this traumatic encounter (Cole, 2004). The obvious benefits for testing color vision as early as possible is to improve a student's academic success in school and decrease the development of poor self-esteem and potential misdiagnosis of a learning disability. It is also advantageous to counsel with students who are 14-21 years of age about possible career choices that are available and prevent career disillusionment. Also, the detection of color vision deficiency can potentially prevent occupational hazards, driving risks, poor labeling, and illness due to decayed or uncooked foods.

1.6 Effects of Abnormal Color Vision on Everyday Tasks

How the human eye perceives color is a contributing component to a range of behavioral functions such as object recognition, visual search, and the evaluation of material properties (Cranwell, Pearce, Loveridge, and Hurlbert, 2015). The world around us provides information contained in light and spectral composition. However, for those individuals with abnormal color vision, visualization of the color-coded world around them significantly compromises their ability to perform both personal and professional tasks (Oliveria, 2014). In 1794, John Dalton gave a prominent speech explaining his own color vision deficiency. The information for this lecture was methodically gathered through his own observations of the presentation of color and lighting conditions, thus came to be termed "Daltonism". Consequently, from that time going forward about 200

years, researchers concentrated predominately on examining the characteristics of color vision deficiency through the use of psychophysical and neuropsychophysical methods as opposed to the experiences of those who have inherited this condition. The thought was that subjective reports from those with color vision deficiency were unreliable and too qualitative in comparison to data procured from "color matching experiments and recording neural responses" (Cole, 2004).

Cole (2004) classifies four types of color tasks in which discrimination and recognition is fundamental. The first color task is comparative, in which a subject is able to discriminately judge various color differences in shade. For instance, certain occupations such as house painting requires color matching as an important part of customer satisfaction. Such occupations could include interior designers, architects, dentists, and industrial and manufacturing. The second color task, referred to as denotative, are those used to establish visual identification, such as, my house is the blue one. The third is aesthetic color tasks which creates an emotional, decorative, and graphic purpose. The fourth color task, termed connotative, is color that is used as a code for relevant information. Color coded information affects various occupations such as commercially related seafaring positions where electronic navigation, radar, and costal/harbor signal lights are critical for the safety of maritime and military vocations.

Other significant occupational positions where color coding is vital for information and identification are (1) railway employment where the recognition of red, green, and yellow signals at considerable distances is essential and affected by poor visibility due to climate conditions, (2) aviation jobs in which navigational aids and the Precision Approach Path Indicator is used for signaling glide paths for landing a plane,

and (3) law enforcement personnel who deal with forensics, identification markers, and the driver safety (Cole, 2004).

Surface color codes transmit detailed knowledge of the environment in various forms. There are two categories that fall under surface color coding, naturally occurring and man-made, that express the importance of testing and counseling early with those who are found to be color vision deficient. Naturally occurring color codes help us to judge everything from the ripeness of fruits, the degree to which our food, especially meats, are adequately cooked (Cole, 2004). Natural surface colors of fruits and foliage present a wide variety of ways where significant information gathered through this code is essentially compromised. Spaulding (2004) described his personal experiences as well as the experiences of other physicians who are color vision deficient during their medical careers. Spaulding (1997) conducted a study with 40 medical practitioners and found that as many as 40 to 60 percent of these practitioners reported difficulty recognizing signs of illness such as cyanosis, jaundice, rashes, and bloody by-products of the body.

Man-made surface color codes are used extensively to support alphanumeric or symbolistic coding. This type of surface coding is usually exhibited in substances such as electrical coding, navigational marks, and in computer displays. Unfortunately, man-made color-coded designs are seldom considered in reference to those with abnormal color vision (Cole, 2004).

Another source to consider is that of chemical color indicators used by physicians, histopathologists, and laboratory technicians. John Dalton, noted earlier for being the first scientist to take academic interest in color vision deficiency, claimed that when

viewing a pair of stockings with a stain, he could scarcely distinguish if the stain was blood or simply dirt.

Since that time, various physicians have reported to have struggled with aspects of their profession due to their color vision deficiency. Such reports include a medical student from 1881 who found it very difficult to use the ophthalmoscope to observe an inflammatory eye. In 1885, author George Wilson reported that interviews with four physicians revealed difficulties in participating in chemistry and in inflamed cheeks and lips. In 1907, Hans Haenal, a clinician, revealed his issue with looking at patients' skin, lips, cheeks, and optic disc. Tocantis and Jones had a publication in 1933 that identified nine of 70 medical students who made mistakes in the observation of stained bacteria, blood cells and colors viewed through a spectroscope. Lastly, in 1951, a physician by the name of Hienz Ahlenstiel, reported that lighter shades of red are difficult to detect or even overlooked and stronger reds appear to him as dark gray (Anthony and Spalding, 2004). Spalding summarized in his study that "medical students are screened at only one university in the United Kingdom and only at a few in the rest of the world" (Spalding, 1999). He also stated in a study involving 40 physicians, some of the most difficult daily tasks were analyzing color coded charts, prints, test-strips, and performing ophthalmoscopic diagnosis.

1.7 Color Testing Instruments

Various color testing instruments have been developed and used for decades. Some of the most popular testing instruments are the Nagel anomaloscope, Farnsworth D-15 color sorting task, Ishihara and Hardy-Rand-Ritter plates, and various computer-generated color tests (e.g., "Color Vision Testing Made Easy" from Waggoner

Diagnostics), designed for children and individuals with disabilities. Many of these have been employed for decades since they are useful in the clinic. More elaborate psychophysical measurements, such as heterochromatic flicker photometry, have also been used, though in a more controlled research environment (Sincich, Sabesan, Tuten, Roorda, and Harmening 2016).

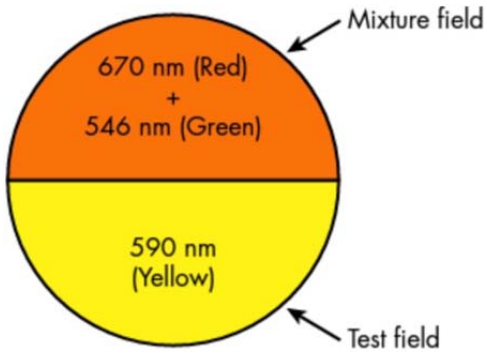


Figure 1.8. Bipartite field of the Nagel anomaloscope. The top field (mixture field) consists of 670 nm and 546 nm wavelength whose intensity balance can be varied by the subject. The luminance of this field, measured for normal trichromacy does not change as the wavelength mixture is varied. The bottom field (test field) consists of 590 nm. This test field's radiance can be adjusted from a very low (dim) to a very high (bright) setting. From Schwartz (2010).

In 1907, the Nagel anomaloscope was developed as a clinical evaluation to recognize those with abnormal color vision, specifically the phenotypic variations in X-linked color disorders (Jägle, Pizer, and Sharpe, 2004). This test instrument uses a small bipartite visual field with two controls; one varies the luminance of yellow light while the other sets the ratio of red to green light (**Figure 1.8**). While this protocol can classify color vision deficiency by assigning a numerical value (when compared to a distribution of normals), other protocols based on error counting can reveal the severity level of a

color deficit.

The problems with the Nagel anomaloscope are that it is bulky and quite expensive. Another issue with the anomaloscope is that even though this is a precision instrument, Jurasevska et al. (2014) state that the correlation between the matching range and the performance of everyday tasks pertaining to color-discrimination is poor.

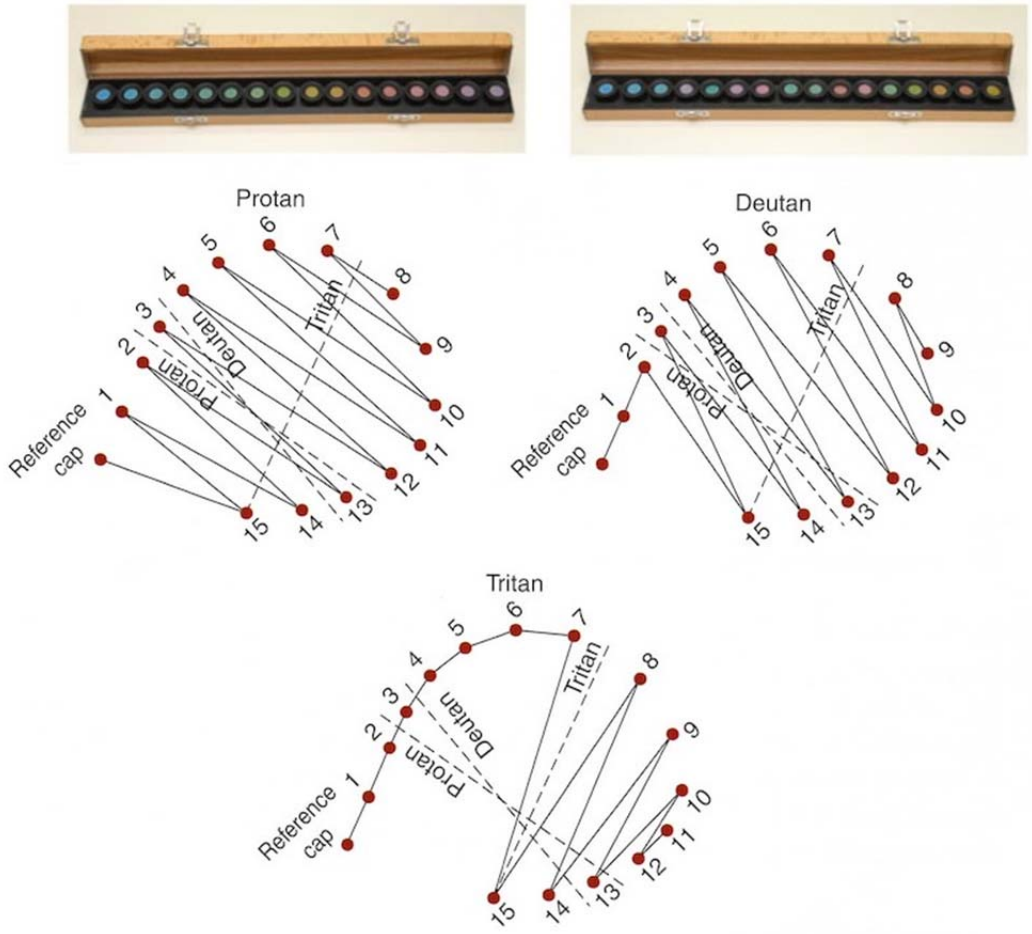


Figure 1.9. Arrangement of the Farnsworth D15 caps in ordered and shuffled format (top). Below are shown the cap orders predicted for protanopic, deuteranopic, and tritanopic anomalies. Diagnosis of an anomaly is made by noting the axis of the crossovers. In a selective color vision loss, the axis corresponds to the protanopic, deuteranopic, or tritanopic color confusion lines, but for nonselective loss there may be no discernable axis. From Schwartz (2010).

The Farnsworth D15 dichotomous color arrangement test was developed in 1947 by Commander Dean Farnsworth for use in the Navy Laboratory to identify those with color vision deficiency who still having the ability to sufficiently discriminate colors in everyday conditions (mild color vision deficiency). This was to distinguish mildly affected personnel from those who have very poor discrimination and are likely to confuse surface color coding (severe color vision deficiency). The "D" in the test's name represents this intent to dichotomize the mild from the severe forms. The test involves a series of colored "caps" that must be arranged in order by hue under a controlled lighting condition. The Farnsworth D15 (**Figure1.9**) is widely used to test potential candidates for occupational positions where color vision is crucial to complete the assigned task with accuracy and safety (Cole and Orenstein, 2003). However, when Dain, Atchison, and Hovis (2019) conducted research on the Farnsworth D15 arrangement test, they subsequently determined that there were a few problems with these panel assessments. Some of the issues mentioned were that sorting tests become soiled over time, working distance, lighting irregularities, practicing to pass or negative malingering, and time allowance, all of which can compromise the utility of the test.

Pseudoisochromatic plates are yet another test for assessing color vision deficiency. The Ishihara test, a color perception test using pseudo-isochromatic plates, was designed in 1917 by Shinobu Ishihara, a professor at the University of Tokyo. The HRR test was developed by LeGrand Hardy, Gertrude Rand, and Catherine Rittler and first published by the American Optical Company in 1955. These plates are essentially printed dots of various colors, brightness, saturation and sizes, typically arranged so that the dots of similar color form a figure (a letter, a numeral or geometrical shape) among a

background of dots of another color or gray. The colors of the figure and the background correspond to the confusion colors of the various types of abnormal color vision. The HRR is currently preferred over the Ishihara pseudoisochromatic plates because research has shown that the HRR has a higher sensitivity to red-green defects and screens for blue-yellow defects as well (Almustanyir, 2020). A dichromat or anomalous trichromat has trouble perceiving these patterns because they are practically undetectable from the background. On the other hand, these portable devices require specific illuminance in order to be valid, which is sometimes neglected especially in clinical use (Kintz,1983).

1.8 Hypothesis

We explored whether we can use readily available materials such as crayons, paint chip samples, and digital projection of synthetic patterns to identify people with deficient red-green color vision. The purpose for these assessments was to potentially provide a reliable alternative to the purchased versions that would still yield accurate results for screening purposes, but with less cost per test. We created four tests for this project: (1.) crayon classification, (2.) crayon color board classification, (3.) paint chip classification, and (4.) red increment test.

Because CVD subjects have poor detection increment thresholds at the long wavelength end of the visible spectrum, a "red test" was developed based on the psychophysical method of constant stimuli. With the method of constant stimuli, a set of stimuli spanning a threshold range is presented in random order over many trials. The stimulus value that elicits a preset level of detection responses. The basis of the red test is that dichromats are insensitive to long wavelength light (Dain and King-Smith, 1981). This is also true of all color vision deficient people, including anomalous trichromats

(York and Loop, 2008). In this research project, a powerpoint presentation was created

The basis for the crayon, colorboard, and paint chip tests is the abnormally perceived saturation around blue-green and some purples. While the Crayola crayon, Crayola color board, and paint chip tests were founded upon basic psychophysical saturation principles, and the results of one prior test subject, the red test is based upon more extensive data (Dain and King-Smith, 1981) (York and Loop, 2008). Furthermore, we presume and will determine if red/green anomalous trichromats will also have difficulty discriminating certain crayon and paint chip colors despite not having a neutral point, specifically those samples what will lie near a dichromat's neutral point (490-500 nm).

CHAPTER 2

Methods

2.1 Subjects

Forty-nine adult human subjects (18-59 years old) were recruited by word of mouth or by research recruitment flyer, with 19 subjects being female and 30 males. Twelve male subjects and two female subjects presented with color vision deficiency, while nineteen males and sixteen females were confirmed to have normal color vision, using the Hardy Rand-Rittler described below. Most subjects were recruited with from Ewing, New Jersey and tested at the College of New Jersey. Written informed consent was obtained according to the guidelines of the Declaration of Helsinki and the University of Alabama, Birmingham's Institutional Review Board. All subjects reported good general and ocular health.

2.2 HRR Standard Color Vision Test

The HRR (Hardy Rand and Rittler) Standard Pseudoisochromatic Test, 4$^{th\ Edition}$ (Baily, Neitz, Tait, & Neitz, 2004) is a color vision assessment that contains 24 test platesfor the identification of the type of defect and diagnosis of color vision deficiency as wellas the extent of the defect. The HRR test provides a quick classification for normal color vision. The HRR is comprised of 4 screening plates for red-green deficiency (**Figure 2.1**)10 plates for classifying protan, deutan and tritan, and 10 plates to classify the deficit

categories as mild, medium or strong (Hardy et al., 1954). We used the HRR as opposed to the Nagel Anomaloscope because the anomaloscope tends to be very expensive and is less portable than the HRR. Subjects were shown the screening plates at a distance of approximately 40 cm while wearing the Gulden C Daylight glasses (lens filters which transform normal incandescent light to "Illuminant C" type illumination taking the place of the Macbeth Easel Lamp). A Bayco SL-300 8.5- inch Clamp Light with Aluminum Reflector mounted on a photography tripod with a 100 watt incandescent light bulb (the

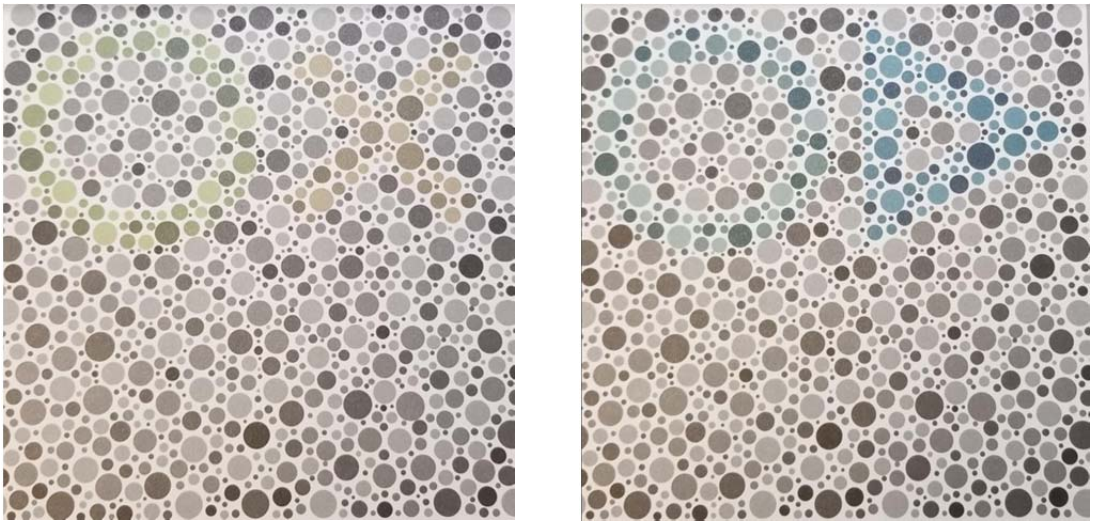


Figure 2.1. Images of the Hardy Rand-Rittler pseudoisochromatic plates 5 and 8.

room was otherwise dark) was angled above the subjectat approximately 60 cm distance from lamp to HRR plates (Moreland and Westland, 2001). All subjects were tested with the HRR pseudoisochromatic plates and scored according to the manufacturer's test instructions. The first four plates presented were demonstration plates made up of colored figures (O, X and Δ) and are not scored.

The subject was asked, "how many figures do you see and what are they?" (Neitz and Bailey, 2002). Once the subject gave an answer, they were asked to trace the symbols with the brush that is provided with the test booklet. Then, subjects were informed that the test is made of the three figures seen in the test plates "with two, one or none on a page and some will be harder to see" (Neitz and Bailey, 2002). Screening plates 5-10 were shown to the subject with the following questions, "How many-colored symbols do you see? What are they? Where are they?", and a response was required (Neitz and Bailey, 2002). The subject's answers were recorded on a datasheet by an experimenter drawing the symbol in the quadrant seen and a check mark was placed beside the box forthe correct answer. Screening plates 6-10 were then shown to the subject allowing only 3second inspection. Once screening plates were completed, a brief review was made to determine if the subject had normal color vision or more testing was required. Plates 5-6 were used to determine if the subject had a blue- yellow defect and plates 7-10 determined a red-green defect in color vision. If the subject incorrectly answered plates 5-6, they were required to continue with plates 21-24. If the subject

incorrectly answeredplates 7-10, they were required to continue with plates 11- 20. If there were incorrect answers in both 5-6 and 7-10 plates, the subject was required to complete all remaining plates (Nietz and Bailey, 2002).

2.3 Crayola Crayon Test

The purpose for the crayon, crayon-based and paint chips tests was to potentially develop a valid, reliable screening tool to assess groups of subjects as opposed to one at a time. The Crayola crayon test is a prototype test we developed using 17 colors from a Crayola 64 pack of crayons. These colors were determined to be the neutral point series by distinguishing colors that looked blue-green or purple which is the general appearance of the colors at the two ends of a protanope's or deuteranope's confusion line passing through white. Dichromats have a point on the spectrum which is the neutral point where one wavelength appears achromatic, which will look like a shade of gray depending on intensity, while the other end of the confusion line ends at a non-spectral purple and also looks achromatic.

The Crayola crayon test was presented to every subject using the 17 chosen crayons in their paper wrap with the name covered by a label with a number (**Figure 2.2**). The names of the chromatic crayons used were *Mauvelous, Robin Egg Blue, Purple, Cerulean, Orchid, Red Violet, Lavender, Purple Mountain Majesty, Turquoise Blue, Blue-Green, Sea green, Wisteria,* while the achromatic crayons were *Black, White,*

Figure 2.2 (Left) Trichromat-selected neutral point series crayons used in the crayon test. (Right) Achromatic grays, white, and black used in the crayon test.

Timberwolf, Silver, and *Gray*. The testing room was illuminated with 18 fluorescent lights and 4 full length windows of approximately 2.44 meters in length and 1.2 meters wide with shades partially open. Ambient light entering from outdoors was not controlled in any other aspect for the crayon, color board, or paint chips test. Prior examination with this method provided preliminary information from a protanope who was presented with various series of colors from the Crayola 64 classic crayons results were as follows:

- Red, Green, Yellow series: When ask to sort 18 crayons (different from those listed above) into piles of "reddish" greenish" or "yellowish" the protanope subject did so quickly and perfectly.
- Neutral point series (**Figure 2.2**): When ask to sort 17, trichromat selected, crayons that appeared '*purple*,' '*blue-green*,' or achromatic *black*, *white*, and *gray* into piles of "colored" and "not colored" he did so quickly and almost perfectly. He put *Mauvelous* and *Robin Egg Blue* in the "not

colored" pile.

Thus, this protanope confused "not colored" (*black, white, grays*) with two particular "colored": *Mauvelous* (purplish) and *Robin Egg Blue* (blue-greenish) presumably due to these colors falling on his achromatic confusion line.

The participant answer sheet (**Figure 2.3**) was a 2-column chart with a title above each column. The title above the first column was "*Black, White, Grays*" and the title above the second column was "*Not Black, White, or Grays."* Seventeen rows were created to accommodate the number of crayons used for this test to allow the participants to mark their choice clearly for how they sorted the crayons.

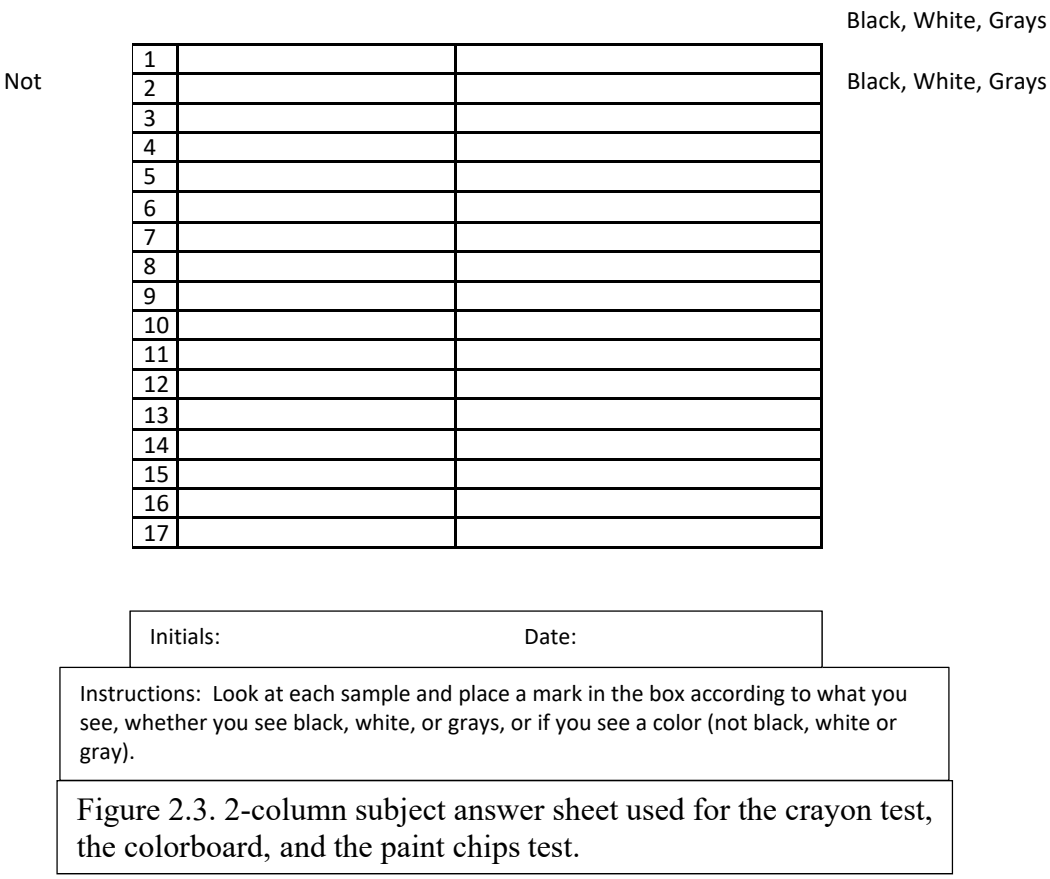


Figure 2.3. 2-column subject answer sheet used for the crayon test, the colorboard, and the paint chips test.

This answer sheet was used for the Crayola crayon test, the color board test, and the paint chips test. An answer key was made using an answersheet provided to participants and corresponded with the numbered crayons.

Analysis of each participant's data was done with Excel (Microsoft). For each participant, a two-column spread sheet was created with the first column labeled "*Not black, white, grays*" with a list of all the colored samples in one column while the second column labeled "*Black, white, grays*" listing the achromatic samples. If a participant marked an answer incorrectly an arrow was drawn (using the draw tap on the tool bar) toward the opposite column indicating that the participant saw that sample as being achromatic. An error rate was calculated for each subject using the number of correct answers by the total number of crayons used. An average error rate was calculated for the normal and color-deficient populations, as identified by the HRR test outcomes.

2.4 Color Board Test

We developed another version of the crayon test by using 17 circles (1.27 cm in diameter) which had been colored in by each crayon from the list provided in the Crayola crayon test (**Figure 2.4**). The subjects were asked to judge each circle as to color (not black, white, or gray) vs. non-color (*black, white*, or *gray*) and recorded their answers on an answer sheet. This experiment was used to determine if the color board color dots are equivalent to the results of judging the crayons themselves. Analysis of each participant's data was done with Excel (Microsoft). For each participant, a two-column spread sheet was created with the first column labeled 'color' and a list of all the trichromatic colors in that column while the second column labeled 'not color listed the achromatic colors. If a participant marked an answer incorrectly, an arrow was drawn

using the draw tab on the tool bar in Excel to the opposite column indicating that the participant saw that color as being achromatic. An error rate was calculated to each subject using the number of answers correct by the total number of colors used. An average error rate was calculated for the normal and color-deficient populations, as defined by the HRR test outcomes.

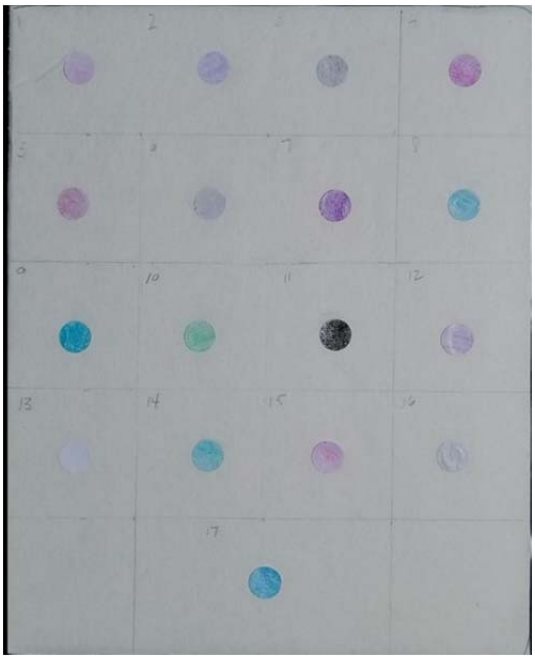


Figure 2.4. Color board test using both trichromatic and achromatic colors. These colors are as follows:

1. Orchid *2. Purple Mt. Majesty*
3. Gray *4 . Red violet*
5. Mauvelous *6. Timberwolf*
7. Purple *8. Blue-green*
9. Turquoise Blue *10. Sea Green*
11. Black *12. Wisteria*
13. White *14. Robin egg blue*
15. Lavender *16. Silver*
17. Cerulean

2.5 Paint Chip Test

We developed a sorting task with paint chips presented to subjects as 17 individual circles (1.27 cm in diameter) cut from each color sample and placed on a 20.3 cm × 25.4 cm foam poster board (**Figure 2.5**). The color scheme was adopted from a Benjamin Moore retail paint store and comprised of samples matching those of the crayon test as closely as possible by two normal trichromats. The colors chosen were *Pink Taffy, Blue Orchid, Pilgrim Haze, Rhododendron, Bayberry, Metallic Silver, Mystical Grape, Fairy Tale Blue, Caribbean Blue Water, Bud Green, Black, Amethyst*

Cream, Super White, Baby Boy Blue,Passion Pink, Laguna Blue, Gull Wing Gray. The subjects judged each paint chip circle to have some color (not *black, white or gray* or no color (black, white, or gray) and recorded their response on a score sheet similar to the Crayola crayon board test score sheet (**Figure 2.3**). An answer key was made using an answer sheet provided to participants and corresponded with the numbered circles. Analysis of participant data was done with Excel (Microsoft). For each participant, a two-column spread sheet was created with the first column labeled 'color' and a list of all trichromatic color in one column while the second column labeled 'not color' listing the achromatic colors. If participants marked an answer incorrectly an arrow was drawn using the draw tab on the tool bar in Excel to the opposite column indicating that the participant saw that color as being achromatic. An error rate was calculated for each subject using the number of answers correct by the total number of colors used. An average error rate was calculated for the normal and color deficient populations, as defined by the HRR test outcomes.

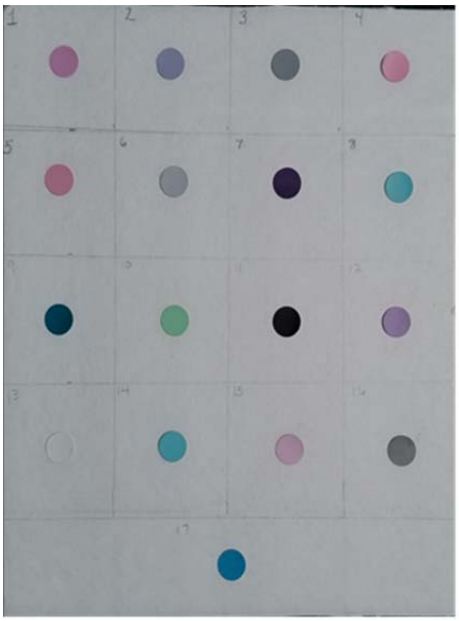


Figure 2.5. The color scheme for the paint chips test was adopted from a Benjamin Moore retail paint store and comprised of samples matching those of the crayon test as closely as possible by normal trichromats.

1. *Pink Taffy*
2. *Blue Orchid*
3. *Pilgrim Haze*
4. *Rhododendron*
5. *Bayberry*
6. *Metallic Silver*
7. *Mystical Grape*
8. *Fairy Tale Blue*
9. *Caribbean Blue Water*
10. *Bud Green*
11. *Black*
12. *Amethyst Cream*
13. *Super White*
14. *Baby Boy Blue*
15. *Passion Pink*
16. *Gull Wing Gray*
17. *Laguna Blue*

2.6 *Red Increment Test*

We developed this test based on York and Loop (2008), who used a red-light increment threshold where a red light with adjustable intensity was added to a bright white background (150 cd/m^2) in order to measure the intensity increment required for a red spot to be detected. The intention of this portion of the research was to determine if Microsoft Powerpoint slides could be used to measure the sensitivity to a red spot increment on a gray background containing random luminance noise.

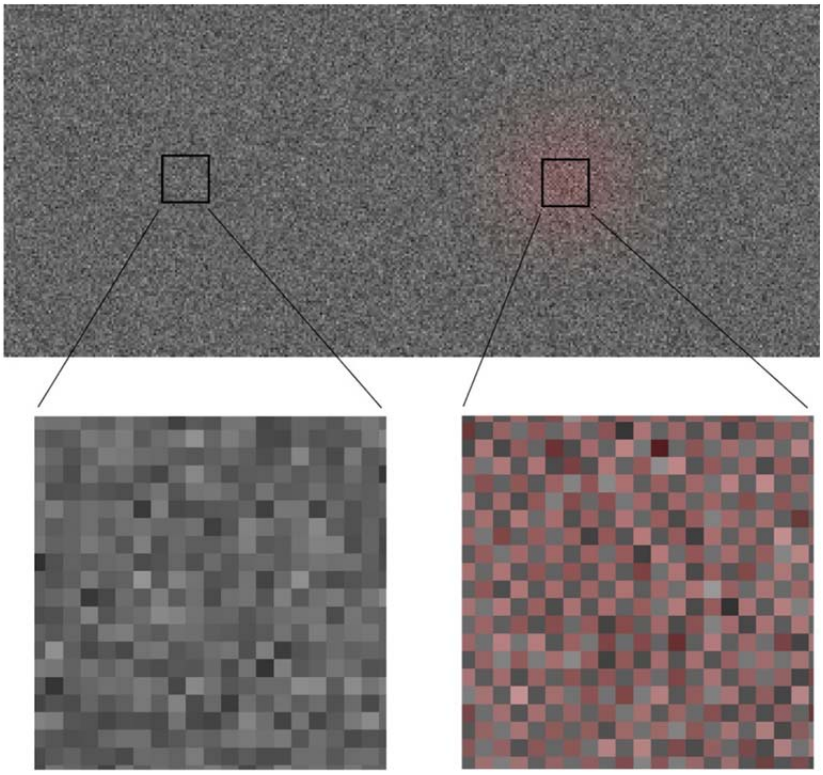


Figure 2.6. The intensity of the red gun was increased in the form of a two-dimensional Gaussian, with a half-width at half-height of 23 pixels. The red increment Gaussian was dithered in every other pixel, as shown in the panels, in order to achieve lower effective contrast for the stimuli.

The stimuli were created in MATLAB (MathWorks, Natick, MA) and then imported into PowerPoint for display. For each trial, a rectangle of grayscale Gaussian noise on a horizontal panel of 750 × 150 pixels with a mean gun level = 100 (on a 0-255 grayscale) and a standard deviation of ±18) was centered in the middle of each slide. At one of 5 possible locations (labeled A-E), the intensity of the red gun was increased in the form of a 2-dimensional Gaussian, with a half-width at half-height of 23 pixels (**Figure 2.6**). The red increment Gaussian was dithered in every other pixel in order to achieve lower effective contrast for the stimuli. The height of the red increment at the center of the Gaussian was scaled logarithmically between 8 and 60 red gun values above the mean background level across 5 different increment levels with subtending 2 degrees. This yielded five increments that were 8, 13, 22, 36, and 60 red gun values added to the noise background. Consequently, the increments at peak for the red values averaged 108, 113, 122, 136, and 160.

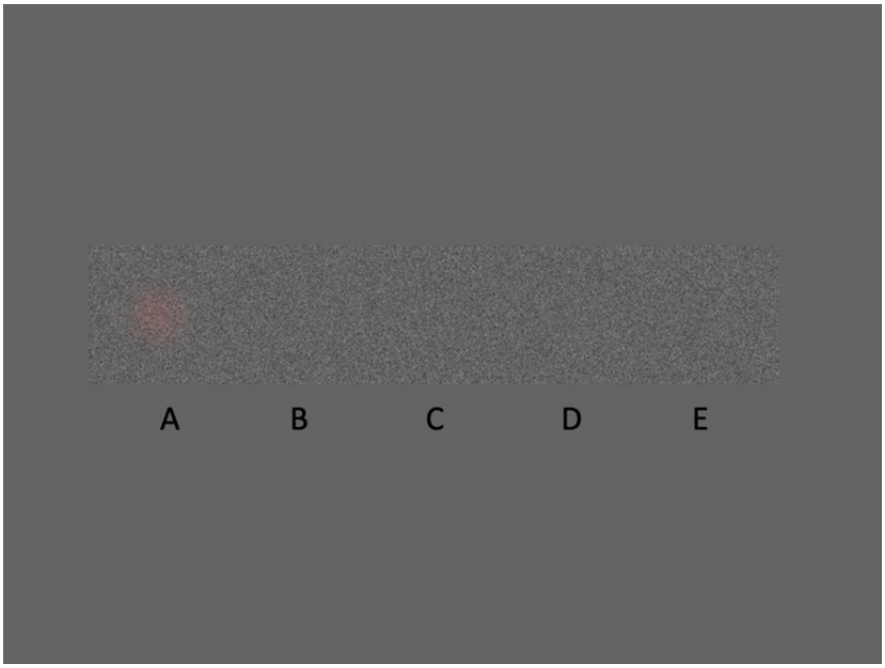


Figure 2.7. Example PowerPoint slide containing a rectangle of Gaussian noise and possible positions which were indicated by letter (A, B, C, D, E) and the increment stimulus (~2 deg in width) which was located above one of the letters.

An example trial is shown in (**Figure 2.7**), with a stimulus located in position A. The trials were grouped in blocks of 5, with each of the 5 intensity levels presented consistently in descending order, but the position of the red increment was randomized within each block of 5 trials. There were 10 repetitions of each stimulus intensity, resulting in 50 trials for the whole experiment for each subject. The task was forced choice; if the subject could not see which location contained a red spot, they were instructed to guess. The trials were paced by the subject. When the subject had made their selection, they simply raised their hand to indicate to the investigator to advance to the next slide.

These slides were displayed on an Epson Bright Link Pro Whiteboard with the use of the BrightLink Pro wall mount projector (**Figure 2.8**). The whiteboard has an interactive area of 70 × 100 inches. The projector's illumination technology is Laser Diode with an aspect ratio of 16:10, 1920 × 1200-pixel array, and a stated contrast ratio of up to 16,000:1 with a white light output of up to 4400 lumens.

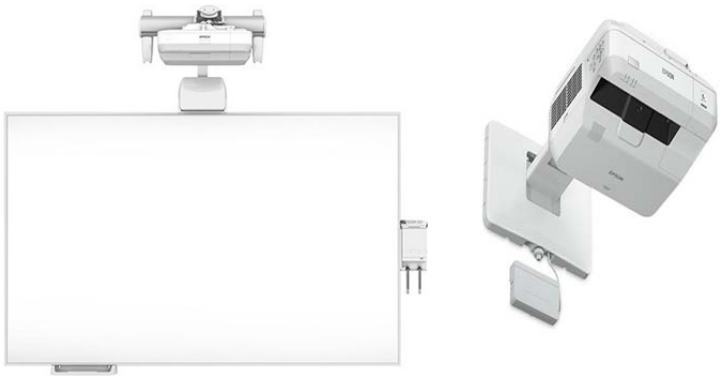


Figure 2.8. An Epson Bright Link Pro Whiteboard with the use of the Bright LinkPro wall mount projector. The whiteboard has an interactive area of 70 × 100 inches. The projector's illumination technology is Laser Diode with an aspect ratio of 16:10, 1920 × 1200-pixel array.

	A	B	C	D	E
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					

You will view 50 slides, please put a mark in the box underneath the letter where you see the stimulus. Each slide is numbered to follow the sequence on the answer sheet.

Figure 2.9. Assessment tool for the use of the red test (only 25 of 50 rows are shown). Each participant was instructed to answer by marking the box underneath the letter in which they saw the stimulus. If they were unable to identify a stimulus in the slide, they were instructed to guess.

Subjects viewed each slide at a distance of approximately 6 meters and used an answer sheet to record their decision about the position of the stimulus. All possible positions were indicated by letter (A, B, C, D, E) and the red increment stimulus

(subtending ~2 deg) was located above one of the letters. Subjects used a 5-column and 50-row scoring sheet to note the perceived stimulus position (**Figure 2.9**).

2.7 Light Level Measurements

The ambient illumination for the Hardy Rand-Rittler, crayon test, color board test, paint chips test, and red test were measured with a Sekonic Light Meter L-858d-U with 1° spot metering and an illuminance range of 0.1 to 2,000,000 lux (**Figure 2.10**). The illuminance for the color board test was 245 lux. Illuminance for the crayon test was 200 lux. The room was illuminated with 18 fluorescent lights and 4 full length windows (north side) of approximately 2.44 meters in length and 1.2 meters wide with the shades partially raised. The external environmental conditions were full sun. For the paint chips test the illuminance at the board was 245 lux with the same environmental conditions as the color board and crayon tests. The Hardy Rand-Rittler illuminance measured at the plates was 740 lux, with light originating from a 100-watt incandescent bulb in a darkened room.

Figure 2.10. Sekonic Light Meter L-858d-U with 1° spot metering, a measuring range of -5 to 22.9(ISO 100), and an illuminance range of 0.1 to 2,000,000 lux.

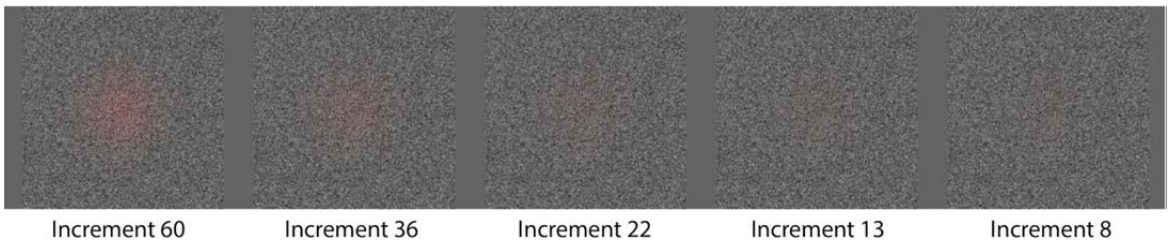


Figure 2.11. The pixel value increments of the five stimuli ranged from 8, 13, 22, 36, and 60. The increments at peak for the red values were 108, 113, 122, 136, and 160. The red increment Gaussian was dithered in every other pixel (see Figure 16) in order to achieve lower effective contrast for the stimuli.

The red increment test was displayed on a uniform gray background. This background luminance was 45 cd/m², while the grayscale Gaussian noise also measured 45 cd/m² over a 1° field. Measurements were taken in the testing room with only one row of florescent lights directly above the subject's seat to allow the subject enough illumination to comfortably make their answers. **Figure 2.11** shows the appearance of the different increments used for the red test. Because of the Gaussian shape and dithering in the stimulus, it was not possible to directly measure the small luminance increments on the Epson Bright Link Pro Whiteboard with the Sekonic light meter. Instead, the values were extrapolated, using large red circular stimuli that filled the meter's measuring window; these stimuli were measured from 0-250 in the red gun only, in steps of 25 (**Figure 2.12**). A smoothed spline fit to these data then allowed extrapolation of the luminance for the stimulus levels used. However, these values only hold for the peak of the Gaussian shape, as the remaining pixel levels in the Gaussian were necessarily lower. To get a more useful estimate of the luminance levels for each stimulus as a whole, the mean of the incremented pixels within the center full width at half maximum of the Gaussian was calculated (**Table 2.1**), showing that the stimulus

increments ranged from 0.75 to 7.69 cd/m^2 in the red gun. These center-most pixels with the highest increments represented 36% of the Gaussian's geometric area. These luminance levels correspond to Weber contrasts of 1.67 to 17.1% on a 45 cd/m^2 background.

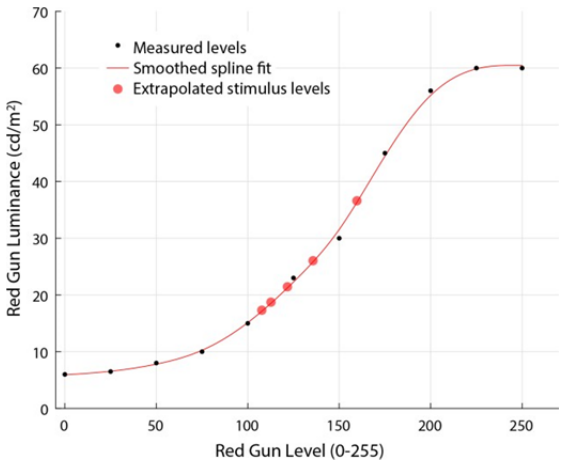


Figure 2.12. Extrapolation of red gun luminance increments. Black dots are luminance values for large noise-free red spots, measured at different gun levels; red line is a smoothed spline fit to these data. Extrapolated luminance values for stimulus levels 108-160 are shown as red circles.

Table 2.1. Stimulus Increments

Red gun pixel increment (0-255)	Peak pixel luminance (cd/m^2)	Gaussian center mean luminance (cd/m^2)	Weber contrast of Gaussian center
8	2.09	0.75	1.67 %
13	3.51	1.26	2.80 %
22	6.22	2.24	4.97 %
36	10.82	3.89	8.65 %
60	21.38	7.69	17.1 %

2.8 Statistical Analysis

All statistics were computed in Matlab version 2016b (Mathworks, Natick, MA), using the available Statistics Toolbox for analysis. Non-parametric statistical tests were used because the data were generally not normally distributed (Kruskal-Wallis test for analysis of variance [ANOVA], and Wilcoxon rank sum test for comparing paired distributions). Subject performance data for the red test were fit with a saturating power law as a function of increment level x:

$$p(\text{correct trials}) = ax^{-b} + c$$

where a = coefficient, b = power exponent, and c = plateau value. Larger values of b produce sharper exponential curves, while c approaches 1 if the proportion p of correct trials gets close to 1.

Receiver operating characteristic (ROC) curve analysis was used to compare the abilities of the HRR plates and the red test in classifying normal and CVD subjects (Hajian-Tilaki, 2013). The binary HRR classification of the 49 subjects (35 normal, 14 CVD), was used as the standard for color vision deficiency classification, against which the sensitivity and specificity of the red test results were computed. Specifically, we calculated the percentage correct detection for the smallest red increment Gaussian center mean luminance (0.75 cd/m^2), since it appeared to offer the best classification ability, as well as the two smallest increments Gaussian center mean luminance (0.75 and 1.26 cd/m^2) to see if the pair improved discriminability. Different red test classification criterion levels, from 10% to 95% of correct detection, were used toclassify the 49 subjects to yield red test classifications of normal and CVD subjects. At each criterion

level, subjects who scored equal to or higher than the criterion were classified as normal

and subjects who scored lower than the criterion were classified as CVD.

One common way to quantify a test's ability to discriminate individuals with and without a disease is to compute the area under the ROC curve. This was done using the trapezoidal function in Matlab. If a test performs poorly, the area under the ROC curve is about 0.5, and will lie close to the unity line. If the test performs well, the area under the ROC curve approaches 1.0. The percentage of correctly detecting the second smallest red increment (1.26 cd/m^2) may also have contributed to CVD discrimination. To take this into account, a compound classifier C was created by a weighted linear combination of the two smallest red increments:

$$C = (P_{0.75} + w\, P_{1.26}) / (1 + w)$$

where $P_{0.75}$ and $P_{1.26}$ were the percentage of corrects of the two smallest red increments and w a weight between 0 and 1. When $w = 1$, the two increments contribute equally to the compound classifier. When $w = 0$, only the 0.75 cd/m^2 increment contributes to the classifier. The linear combination is divided by $(1 + w)$ so that C always has a value between 0 and 1. Different weight values were used to calculate what weighting of the compound classifiers yielded the largest area under the ROC curve, thereby identifying the optimal classifier.

CHAPTER 3
Results

3.1 Color Vision Classification of Subjects

The study population consisted of 49 subjects ranging in age from 19 to 59 years, both male and female, drawn from African American, Caucasian, Hispanic, Puerto Rician, Asian Indian, and Asian/Pacific Islander ethnicities (**Table 3.1**). All subjects had self-reported good ocular health. The Hardy Rand-Rittler (HRR) pseudoisochromatic plates were used to categorize each subject's color vision, in terms of degree (mild, medium, and strong) and cone deficit axis (protanopic, deuteranopic, tritanopic and red-green unclassified). The category "red-green unclassified" refers to the HRR test's inability to distinguish between the type of red-green color classification of protanopia or deuteranopia, if the number of errors is the same in both the protanopic and deuteranopic columns, or if errors were made only in the screening plates.

Of the 49 participants, 35 were determined to have normal color vision and 14 were determined to have a color vision deficiency (CVD). There were 19 males and 16 females in the group of normal color vision participants, while in the group of color vision deficient subjects, 2 were female and 12 were male. The high percentage of CVD among males is to be expected (Deeb, 2005). The median age for the two groups were not different ($p = 0.9$, Wilcoxon rank sum test), with the CVD median age being 30 years old and the normal participants being 29 years old.

Table 3.1. Subject Data

Subject	Test Date	Initials	Gender	Age	Ethnicity	HRR Outcome
1	11/12/2018	AS	F	21	White	Mild Deutan
2	11/14/2018	AM	M	21	White	Mild Protan
3	5/1/2019	DL	M	26	White	Mild Protan
4	6/11/2019	TK	M	50	White	Mild Protan
5	5/8/2019	TH	M	19	White	Mild RG Unclassified
6	11/15/2018	SR	M	37	White	Mild RG Unclassified
7	10/9/2019	DeN	M*	26	Black	Mild RG Unclassified
8	10/31/2018	DH	M	43	White	Mild RG Unclassified
9	10/9/2019	MV	M	26	Black	Mild RG Unclassified
10	1/3/2019	JH	F	42	White	Mild RG Unclassified
11	10/31/2018	MD	M	21	White	Medium Protan
12	10/31/2018	RS	M	39	White	Medium Tritan
13	12/13/2018	JE	M	45	White	Strong Protan
14	4/22/2019	JD	M	34	White	Strong RG Unclassified
15	4/22/2019	SS	M	41	White	Normal
16	4/5/2019	SD	F	20	White	Normal
17	4/18/2019	SG	F	47	White	Normal
18	11/3/2019	VS	F	27	White/Hispanic	Normal
19	7/25/2019	MC	M	54	White	Normal
20	1/3/2019	SF	M	56	White	Normal
21	1/28/2019	JB	M	45	White	Normal
22	3/26/2019	TP	M	30	Latino/Hispanic	Normal
23	12/13/2018	JM	F	34	White	Normal
24	7/2/2018	JC	M	34	White	Normal
25	10/25/2018	AS	M	21	White	Normal
26	11/8/2018	CS	M	22	White	Normal
27	10/31/2018	BF	M	59	White	Normal
28	10/31/2018	KS	F	29	White	Normal
29	10/31/2018	AR	M	33	Puerto Rican/Korean	Normal
30	11/19/2018	MS	F	21	White	Normal
31	11/19/2018	EV	F	20	White	Normal
32	11/20/2018	CS	F	45	White	Normal
33	11/26/2018	MP	M	20	Asian Indian	Normal
34	5/8/2019	JW	F	19	White/African Amer.	Normal
35	5/16/2019	JD	M	36	White	Normal
36	5/29/2019	KH	M	24	White	Normal
37	8/23/2019	TR	M	41	White	Normal
38	9/25/2019	MC	F	22	Asian	Normal
39	9/26/2019	JC	M	22	White	Normal
40	9/26/2019	AW	F	33	Black	Normal
41	10/7/2019	DN	M*	26	Black	Normal
42	10/9/2019	JB	M	21	White	Normal
43	10/21/2019	AP	M	23	Asian Indian	Normal
44	10/24/2019	SG	F	22	White	Normal
45	11/5/2019	CF	F	37	White	Normal
46	11/5/2019	NT	F	32	White	Normal
47	11/5/2019	BD	F	29	White	Normal
48	4/5/2019	PL	F	22	Asian/Pacific Islander	Normal
49	4/5/2019	MR	F	54	Asian/Pacific Islander	Normal

* Sibling fraternal twins

Of those 14 color vision deficient participants, Subjects 1-10 were mild, Subjects 11-12 were medium, and Subjects 13-14 were classified with strong color deficits from the HRR test. Among the mild CVDs, 1 was classified as a deuteranope, 3 were classified as protanopes, 6 were mild red-green unclassified. Among the medium and strong CVDS, 1 was classified as a medium protanope, 1 a strong protanope, 1 a strong red-green unclassified, and 1 medium tritan. The sampling across ethnicity relating to color vision deficiency was too low to draw any conclusions between the two variables.

3.2 Color Identification Test: Crayon Based and Color Board

The first of the 3 assessments given was the crayon test, whereby subjects separated 17 colors into two groups of chromatic and achromatic colors. Among the normal population, 33 of 35 subjects performed the crayon test with 100% accuracy (**Table 3.2**). Subject 15 categorized the chromatic crayon *Blue-green* as achromatic, while Subject 18 confused the achromatic crayon *Silver* as having some color. This suggests that the use of crayons for a color discrimination task may yield some errors in a population otherwise considered to have normal color vision.

Subjects 1-10 were classified as having a mild CVD, and only one of these subjects classified chromatic crayons, *Purple Mt. Majesty* and *Mauvelous*, as achromatic (**Table 3.2**). Although the numbers of subjects were low, it is notable that 10% of the mild CVDs and 6% of the normal subjects made errors with crayon classification, indicating that the crayon test does not discriminate well between mild CVDs and normals.

Table 3.2. Color Identification Errors Data

Subject	HRR outcome	Color Board	Crayons	Paint Chips
1	Mild Deutan	---	---	---
2	Mild Protan	---	---	Metallic Silver
3	Mild Protan	---	---	---
4	Mild Protan	---	---	---
5	Mild RG Unclassified	---	---	---
6	Mild RG Unclassified	---	---	---
7	Mild RG Unclassified	Timberwolf	---	---
8	Mild RG Unclassified	---	---	---
9	Mild RG Unclassified	Black, Wisteria	---	Laguna Blue
10	Mild RG Unclassified	---	Purple Mt. Majesty, Mauvelous	Blue Orchid
11	Medium Protan	---	---	---
12	Medium Tritan	---	---	---
13	Strong Protan	Timberwolf, Silver	Gray, Timberwolf	Pilgrim Haze, Metallic Silver, Gull Wing Gray
14	Strong RG Unclassified	Gray, Timberwolf, Black, Silver	Gray, Timberwolf, Silver	Pilgrim Haze, Metallic Silver, Gull Wing Gray
15	Normal	---	Blue-green	---
16	Normal	Timberwolf	---	---
17	Normal	Silver, Timberwolf	---	---
18	Normal	---	Silver	---
19	Normal	White	---	---
20	Normal	Gray	---	---
21	Normal	Black	---	---
22	Normal	White	---	---
23	Normal	Silver	---	---
24	Normal	---	---	---
25	Normal	---	---	---
26	Normal	---	---	---
27	Normal	---	---	---
28	Normal	---	---	---
29	Normal	---	---	---
30	Normal	---	---	---
31	Normal	---	---	---
32	Normal	---	---	---
33	Normal	---	---	---
34	Normal	---	---	---
35	Normal	---	---	---
36	Normal	---	---	---
37	Normal	---	---	---
38	Normal	---	---	---
39	Normal	---	---	---
40	Normal	---	---	---
41	Normal	---	---	---
42	Normal	---	---	---
43	Normal	---	---	---
44	Normal	---	---	---
45	Normal	---	---	---
46	Normal	---	---	---
47	Normal	---	---	---
48	Normal	---	---	---
49	Normal	---	---	---

--- = No error, Black font = uncolored item, Red font = colored item

Subjects 11 and 12, who were rated to have a medium deficit, made no errors on the crayon test. Only the results of Subjects 13 and 14 who were classified as having a strong color vision deficiency had outcomes on the crayon test that demonstrated relatively high error rates. These subjects confused the achromatic colors *Gray, Timberwolf*, and *Silver* for chromatic colors, meaning they thought all were colors. This error is different from the one mild CVD who classified some chromatic colors as achromatic.

The second crayon-based assessment given to subjects was the color board test where individuals were asked to determine if they saw color or did not see color when looking at color samples made by shading the same 17 crayons on a white background. Of those with normal color vision, Subjects 16-17 and 19-23 made identification errors confusing the achromatic colors *Timberwolf, Silver, White, Gray*, and *Black* with chromatic colors. The higher error rate among normals (20%) for the color board test suggests that the task is more difficult, as might be expected from the desaturated appearance of the crayon sample when drawn on a white surface.

Subjects 1-6 and 8 who were classified as having a mild color vision deficiency were able to identify all colors on the color board test with 100% accuracy. Subject 7 misidentified or confused the achromatic color, *Timberwolf*, for a chromatic color and Subject 9 confused the achromatic color, *Black*, for a chromatic color, but also confused the chromatic color *Wisteria* for an achromatic color. As a group, the mild CVD subjects also had a 20% error rate, like the normal group, suggesting that this test cannot discriminate mild CVD subjects.

Subjects 11 and 12 were classified as having a medium color vision deficiency

per their results on the HRR, but were also able to identify with 100% accuracy the colors presented in the form of the color board test. This is consistent with their performance with the direct crayon classification and suggests that medium CVD subjects are also not well identified by the color board test.

Subjects 13 and 14 had a classification of strong color vision deficiency and made several color identification errors, grouping several of the achromatic colors *Gray, Silver, Black* and *Timberwolf* with other chromatic colors. This behavior was consistent with the crayon sorting task, with the addition of *Black* thought to be a chromatic color for one subject.

3.3 Color Identification Test: Paint Chips

The third test was the paint chips test. This test was designed very similarly to the color board test except paint chip samples were chosen that were similar to the colors used on the color board test. In viewing **Table 3.2**, normal subjects made no errors with the paint chips test, which is quite distinct from the crayon or color board tests. Three of the mild CVD subjects made single errors: Subject 1 confused the achromatic color *Metallic Silver* for a chromatic color, Subject 9 confused the chromatic color *Laguna Blue* for an achromatic color, and Subject 10 confused the chromatic color *Blue Orchid* for an achromatic color. The two medium CVD subjects made no errors, consistent with their performance on the other two-color identification tasks. For the two strong CVD subjects, their errors were relatively high, classifying 3 of the achromatic paint chips as being chromatic (*Pilgrim Haze*, *Metallic Silver*, and *Gull Wing Gray*).

Their performance is consistent with the other two-color identification tasks, where they repeatedly placed achromatic samples in a chromatic category. Thus, the errors seen across the CVD population with the paint chips test essentially mirrored those made withthe crayon-based tests, and as a whole, the crayon-based tests and the paint chips test did not seem to possess the capability of classifying CVD subjects from normal subjects.

3.4 Red Increment Test

The final assessment used to test subjects was based on the ability to detect a red luminance increment amidst grayscale noise, which we called the red test. The red test was a psychophysical measurement that required the subject to make a decision (five-alternative forced choice) that identified the location of a Gaussian shaped red spot stimulus appearing randomly among 5 possible locations, as detailed in the Methods. Five increment values were used, with 10 trials for each increment, and the proportion of correctly identified locations by each subject were calculated for each increment level (**Table 3.3**).

Table 3.3. Red Test Data

Values = proportion of correct trials

Subject	HRR Outcome	Luminance increment (cd/m^2) 7.69	3.89	2.24	1.26	0.75	
1	Mild Deutan	1.0	0.9	0.8	0.8	0.6	
2	Mild Protan	1.0	1.0	0.9	0.4	0.1	
3	Mild Protan	1.0	0.6	0.5	0.3	0.1	
4	Mild Protan	1.0	1.0	1.0	0.9	0.2	
5	Mild RG Unclassified	1.0	0.8	0.8	0.8	0.8	
6	Mild RG Unclassified	0.9	1.0	0.8	0.3	0.2	
7	Mild RG Unclassified	1.0	0.9	1.0	0.8	0.2	
8	Mild RG Unclassified	1.0	1.0	1.0	0.9	0.2	
9	Mild RG Unclassified	1.0	1.0	1.0	0.6	0.5	
10	Mild RG Unclassified	0.9	1.0	0.9	0.3	0.3	
11	Medium Protan	0.8	1.0	0.9	0.9	0.1	
12	Medium Tritan	1.0	1.0	1.0	0.8	0.1	
13	Strong Protan	0.7	0.8	0.7	0.6	0.3	
14	Strong RG Unclassified	0.3	0.4	0.3	0.2	0.2	
		0.900	0.886	0.829	0.614	0.279	Mean
		0.196	0.183	0.209	0.263	0.212	SD
		1.0	1.0	0.9	0.7	0.2	Median value
15	Normal	1.0	1.0	0.9	1.0	1.0	
16	Normal	1.0	1.0	1.0	1.0	0.6	
17	Normal	1.0	1.0	1.0	1.0	0.7	
18	Normal	1.0	1.0	1.0	0.8	0.7	
19	Normal	0.9	1.0	1.0	1.0	0.2	
20	Normal	1.0	1.0	0.8	0.6	0.3	
21	Normal	1.0	1.0	1.0	1.0	1.0	
22	Normal	0.9	1.0	1.0	1.0	1.0	
23	Normal	1.0	0.9	1.0	0.6	0.3	
24	Normal	1.0	0.8	0.8	0.7	0.3	
25	Normal	0.9	1.0	0.9	1.0	0.8	
26	Normal	1.0	1.0	1.0	1.0	0.9	
27	Normal	1.0	1.0	1.0	1.0	0.2	
28	Normal	1.0	1.0	1.0	0.8	0.3	
29	Normal	1.0	1.0	1.0	0.5	0.1	
30	Normal	1.0	1.0	1.0	1.0	0.5	
31	Normal	1.0	1.0	1.0	1.0	0.9	
32	Normal	1.0	1.0	1.0	1.0	0.2	
33	Normal	1.0	0.9	0.8	0.8	0.8	
34	Normal	1.0	1.0	1.0	1.0	0.7	
35	Normal	0.9	1.0	0.9	1.0	0.6	
36	Normal	1.0	1.0	1.0	1.0	1.0	
37	Normal	1.0	1.0	1.0	1.0	1.0	
38	Normal	1.0	1.0	1.0	1.0	1.0	
39	Normal	1.0	1.0	1.0	1.0	1.0	
40	Normal	1.0	1.0	1.0	1.0	0.9	
41	Normal	0.9	1.0	1.0	1.0	0.8	
42	Normal	1.0	1.0	0.9	1.0	0.7	
43	Normal	1.0	1.0	1.0	1.0	0.8	
44	Normal	1.0	1.0	1.0	1.0	1.0	
45	Normal	1.0	1.0	1.0	1.0	1.0	
46	Normal	1.0	1.0	1.0	1.0	0.4	
47	Normal	1.0	1.0	1.0	1.0	0.9	
48	Normal	1.0	1.0	1.0	1.0	0.7	
49	Normal	1.0	1.0	1.0	0.9	0.6	
		0.986	0.989	0.971	0.934	0.683	Mean
		0.036	0.040	0.062	0.137	0.293	SD
		1.0	1.0	1.0	1.0	0.7	Median value

The group identified as normal in the HRR classification (Subjects 15-49) performed fairly well in this task. For the 4 largest increment levels, the mean proportion of correct trials exceeded 0.93 (**Figure 3.1**), indicating less than 7% errors, with no difference in performance (ANOVA, $p = 0.27$). The smallest increment level, 0.75

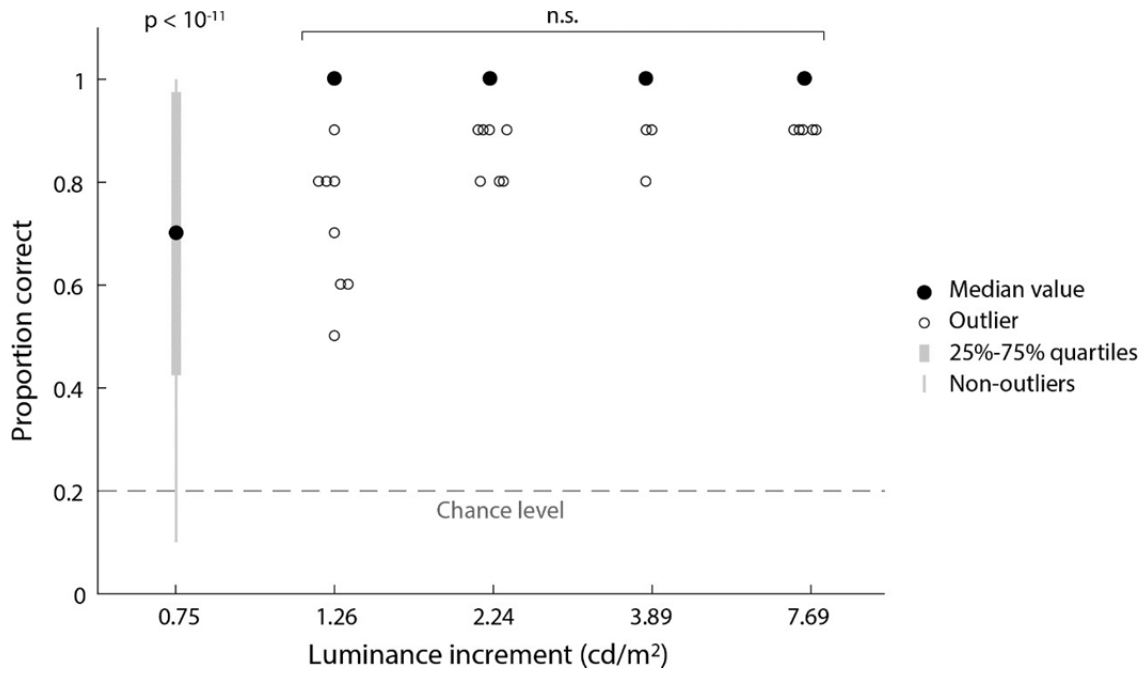


Figure 3.1. Population data of normal subject performance on the red test. Data represent the proportion of correctly detected locations of the red spot for 10 trials at each luminance increment level. n.s. = not significant.

cd/m^2, had a performance of correctly identifying the stimulus location of 0.68±0.29 (mean ±1 SD), which was lower than the 3 largest increment trials (ANOVA, $p = 4.7\times10^{-12}$). Thus, the smallest increment was more difficult to detect for normals, but was still suprathreshold, as the performance was not near the 0.2 chance level.

The group identified as CVD by the HRR classification (Subjects 1-14)

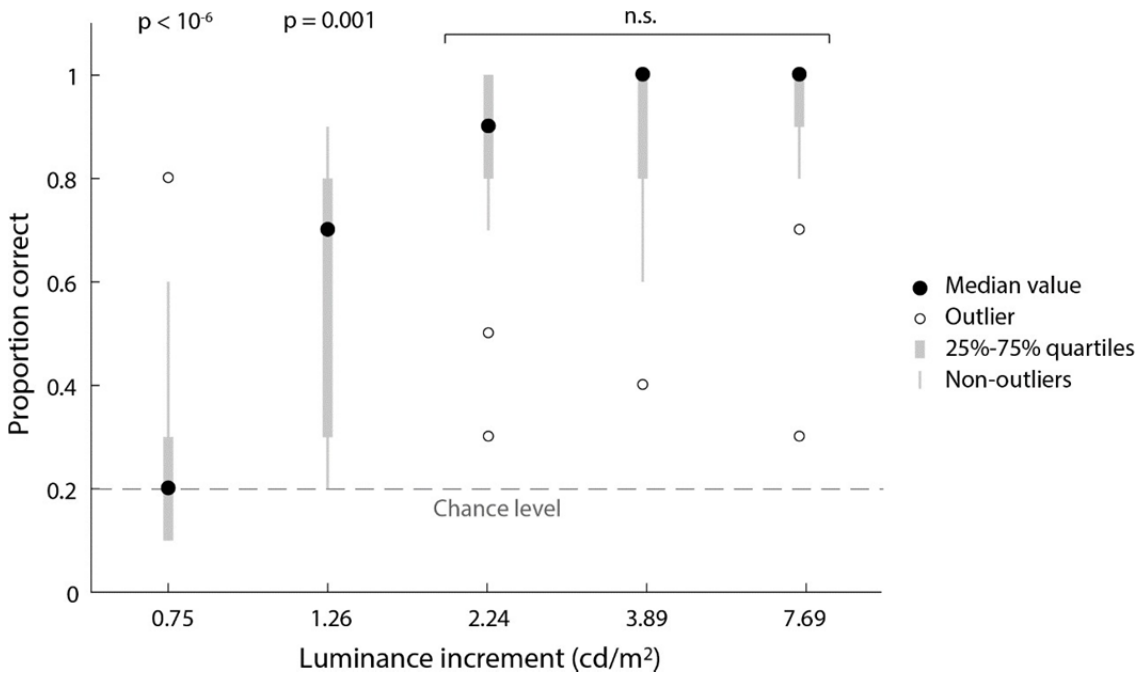


Figure 3.2. Population data of CVD subject performance on the red test. Data represent the proportion of correctly detected locations of the red spot for 10 trials at each luminance increment level. n.s. = not significant.

performed less well on average for the 3 largest increment levels, ranging from 0.90 to 0.83, indicating up to 17% errors (**Figure 3.2**), but performance on these 3 levels were not different from one another (ANOVA, $p = 0.34$). The next-to-smallest increment, 1.26 cd/m^2, yielded a performance of correctly identifying the stimulus location of 0.61±0.26 (mean ±1 SD), which was lower than the larger increments trials (ANOVA, $p = 0.001$), but still suprathreshold. The smallest increment level, 0.75 cd/m^2, lead to chance level performance of 0.28±0.21 (mean ±1 SD), well below the larger increment trials (ANOVA, $p = 1.2\times10^{-7}$). This lowest increment performance suggests that CVD subjects were guessing and could not detect the presence of the smallest red increment.

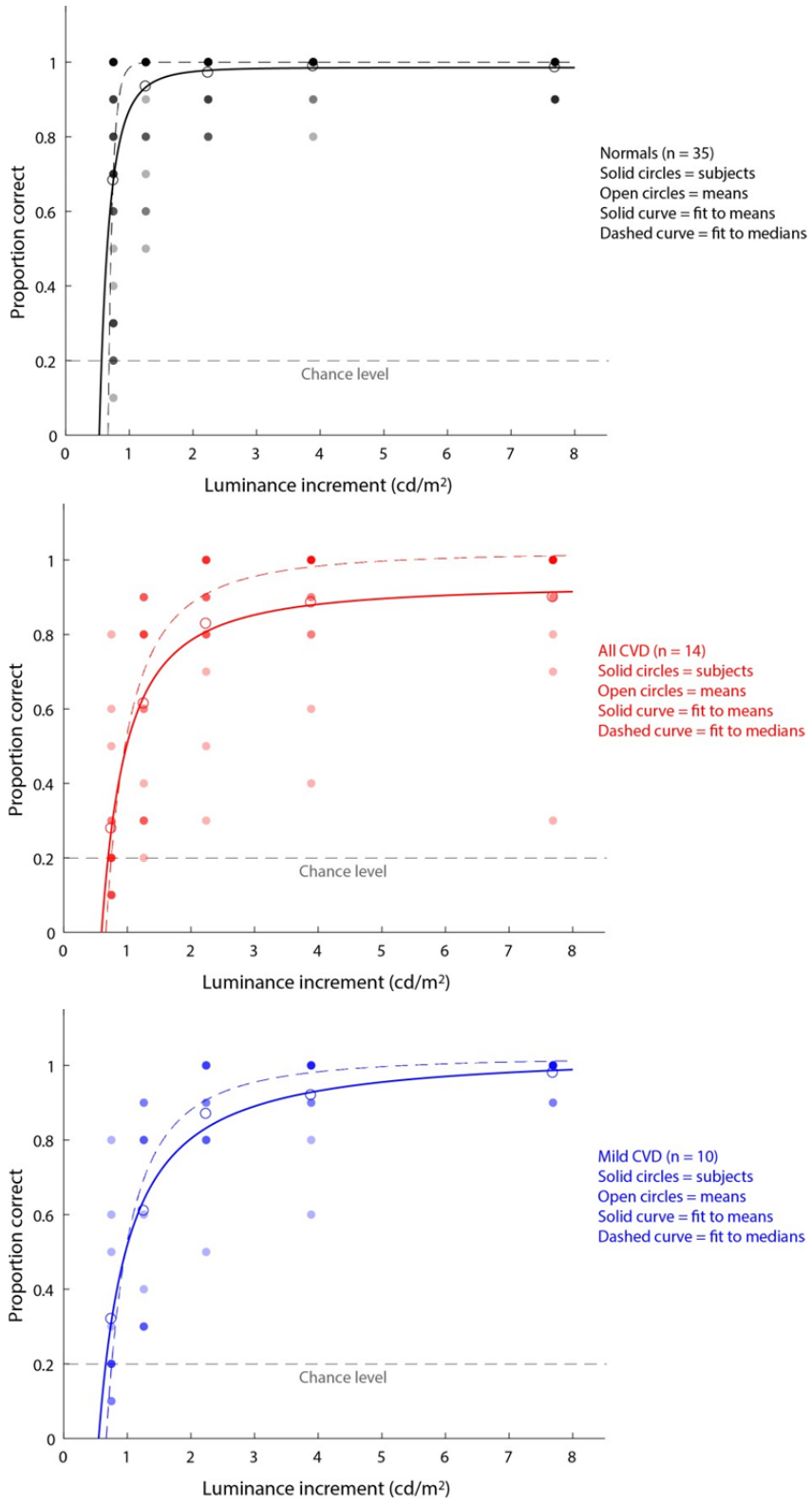


Figure 3.3. Power law fits to data from normal (black), all CVD (red), and mild CVD (blue) subjects. Each subject's data is plotted as a transparent circle; where data points overlap, the color density is higher.

To more easily compare different groups of subjects, we produced power law fits to the data to generate curves representing the median and mean values of the normal, all CVD, and mild CVD subjects (**Figure 3.3**). The 4 medium and strong CVD subjects were too few in number to create meaningful fits to their data. The mean curves were fit from the subject data directly and lie close to the mean values for each increment (which are plotted as open circles in each of the graphs for reference).

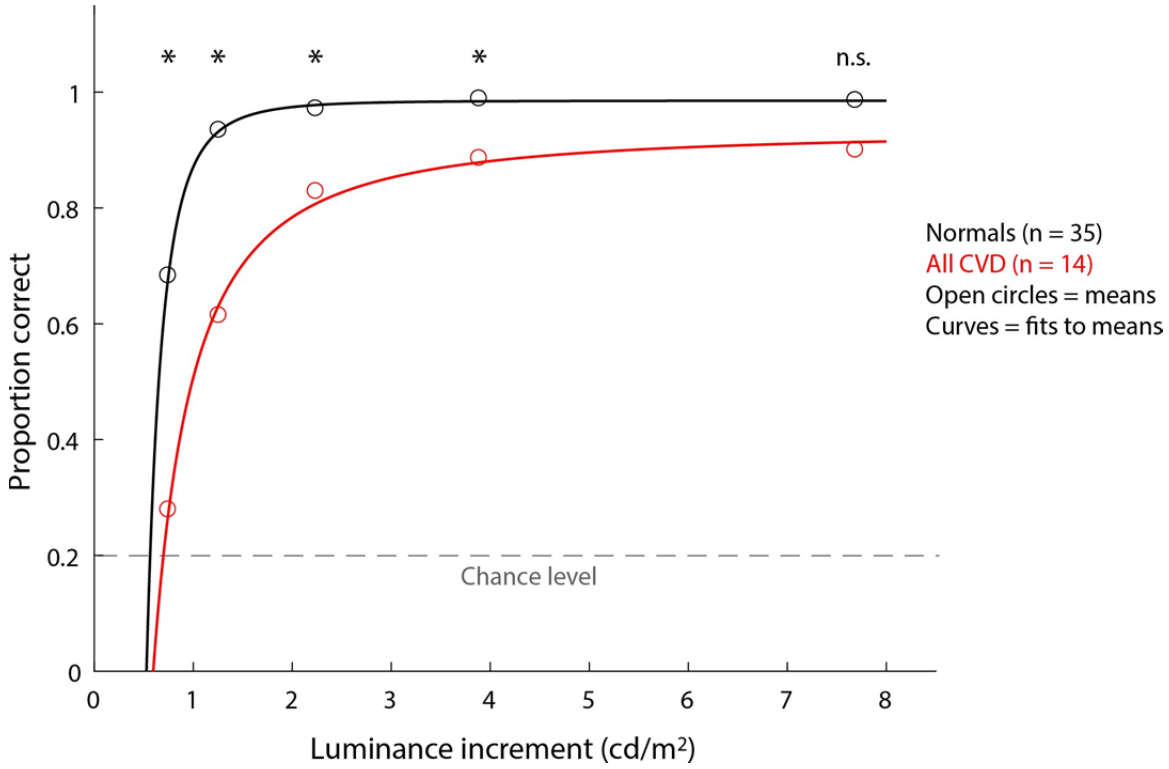


Figure 3.4. Comparison of normal versus CVD subject performance on the red test, plotted as mean values with curve fits. Significance levels for each increment are as follows: 0.75 cd/m^2, p = 7.83×10^{-5}; 1.26 cd/m^2, p = 1.89×10^{-6}; 2.24 cd/m^2, p = 0.0014; 3.89 cd/m^2, p = 0.0041; 7.69 cd/m^2, p = 0.061 (n.s. = not significant).

To directly compare the results between the normal and CVD subjects, the mean data and their fits for each group were plotted together (**Figure 3.4**). There was a difference in performance between the two groups across the 4 smaller increment levels.

(See **Figure 3.4** legend for p values, Wilcoxon rank sum test). These data suggest that performance across the highest set of increments was worse for CVDs than for normals.

In this study, the participants were divided into 4 groups according to their classification with the HRR. To determine if the differences seen in CVD population

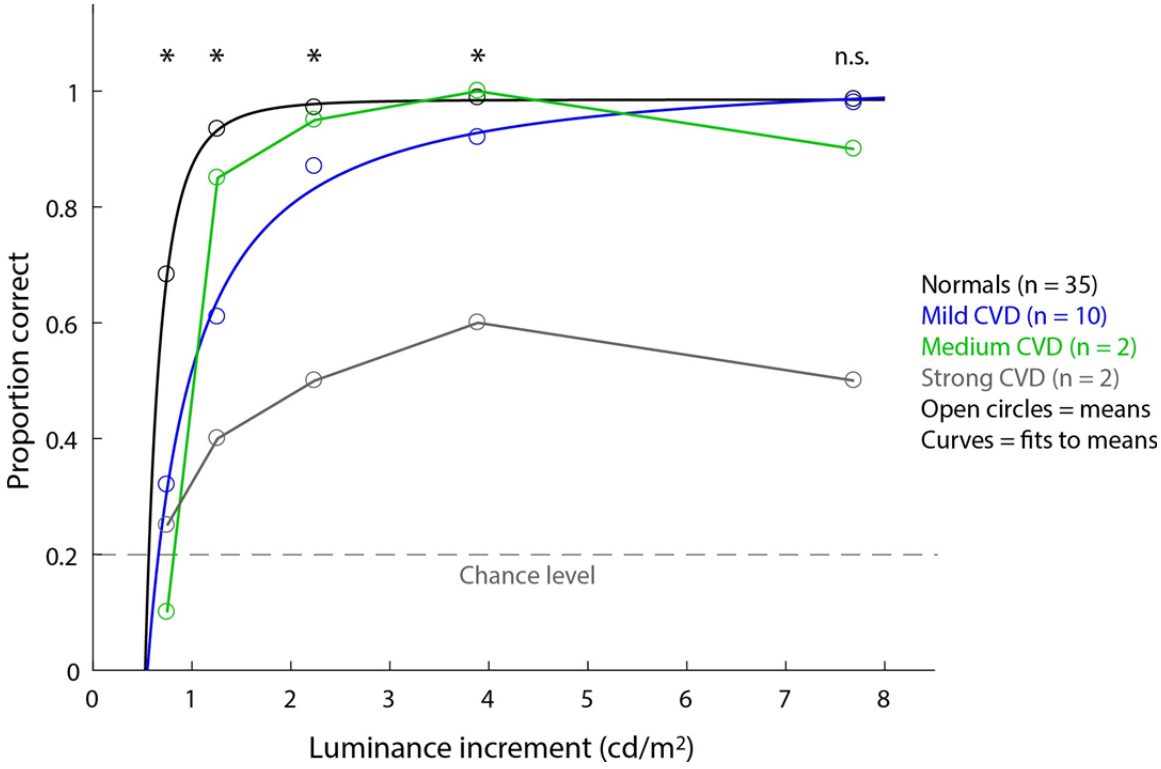


Figure 3.5. Comparison of normal versus CVD subject performance on the red test, separated by CVD category (mild = blue, medium = green, strong = gray). The proportion of correct trials generally diminished with increasing color deficiency, although the low number for the most affected subjects (drawn with connecting lines) make it impractical to draw any firm conclusions. Significance levels between normals and mild CVDs for each increment are as follows: 0.75 cd/m^2, p = 0.0013; 1.26 cd/m^2, p = 1.33×10^{-5}; 2.24 cd/m^2, p = 0.0093; 3.89 cd/m^2, p = 0.015; 7.69 cd/m^2, p = 0.68 (n.s. = not significant).

versus normals was driven by the severity of the CVD deficit, we analyzed the data by subgroup (**Figure 3.5**). When the normals were compared to the mild CVD group alone (n = 10), there was a significant difference in performance across the 4 smaller

increment levels (see **Figure 3.5** legend for *p* values, Wilcoxon rank sum test). The CVDs with medium and strong deficits (each n = 2) often had lower performance than the other groups, as seen in **Figure 3.5**, but had too few subjects to conduct meaningful statistics. All CVD subjects appeared to be performing at or near chance for the lowest increment trials, and the strongest CVDs appeared to be the worst performers at all increment levels. Despite these limitations, the results yielded three outcomes: (1) the red test can identify all degrees—mild, medium, and strong—of CVD classified by the HRR, (2) at the lowest increment, 0.75 cd/m^2, all CVD subjects perform at chance levels, and (3) those subjects with the greatest color vision deficiency are not the source for the CVD difference in the population as a whole when compared to normals.

Receiver operating characteristic (ROC) curve analysis was performed to determine how well the red test performs (percentage correct detection) as a classifier of CVDs compared to the HRR test. For example, when the smallest red increment was used as the classifier and 60% correct was used as the classification criterion, 2 of the 14 HRR CVD subjects were classified as normal and 10 HRR normal subjects were classified as CVD. A 2×2 contingency table for this 60% percent correct criterion yields the following table:

	Red Test: Normal	*Red Test: CVD*	
HRR: Normal	True negative = 25	False positive = 10	HRR (-) = 35
HRR: CVD	False negative = 2	True positive = 12	HRR (+) = 14
	Red test (-) = 27	Red test (+) = 22	Total = 49

From the above table, the true positive rate (sensitivity) and the true negative rate (specificity) of the smallest red increment to detect CVD were:

Sensitivity = True positive / HRR positive = 12/14 = 0.857
Specificity = True negative / HRR negative = 25/35 = 0.714

Each red test classification criteria produces a 2×2 contingency table and a pair of sensitivity and specificity values. Using percent correct criteria ranging from 10% to 95% (**Table 3.4**), a ROC curve was computed with the true positive rates (sensitivities) as the vertical coordinates and the false positive rates (1-specificities) as the horizontal coordinates (**Figure 3.6**).

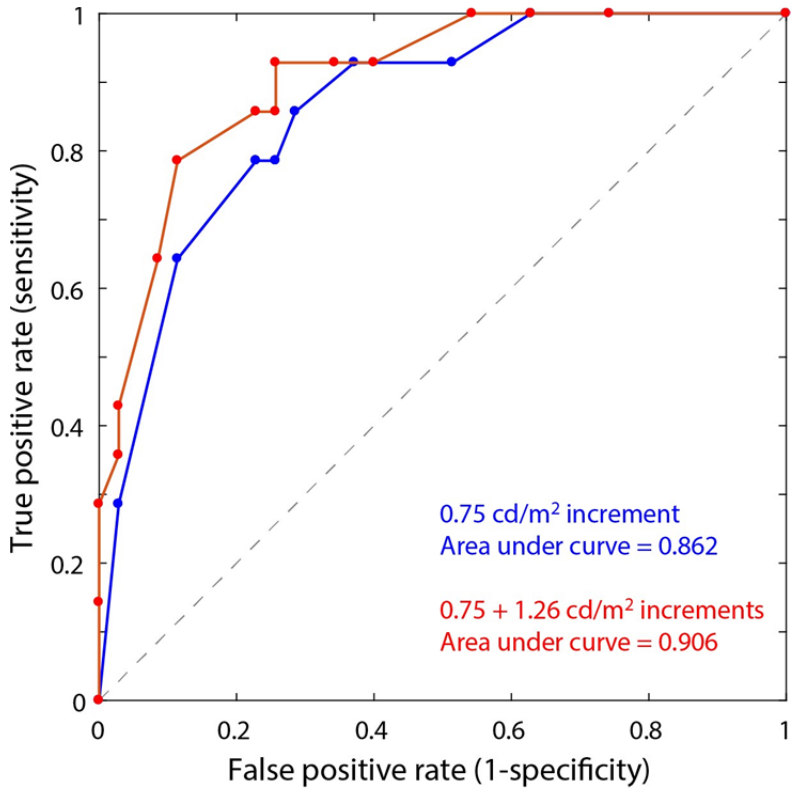


Figure 3.6. ROC curves for the smallest increment (blue) and a weighted combination of the two smallest increments (red). Dashed line is the lower theoretical limit where test data would yield no classification value (area under curve = 0.5).

If this red test classifier separates normal and CVD subjects by chance, the ROC curve would lie close to the dashed unity line. If the red test classifier separates CVDs from normals as accurately as the HRR, its ROC curve should distance itself from the dashed line and hug the upper left corner of the plot. The blue curve in **Figure 3.6** is the ROC curve when the percentage of correct values of the smallest red increment are used as the classifier to classify the 49 subjects. The area under the ROC curve is a measure of a test's ability to discriminate subjects, and for the smallest increment this area is 0.862, which indicates excellent discrimination (Hosmer et al 2013). The optimal percent correct criterion appears to be near 60% based on common metrics used to select the best criterion from ROC curves (**Table 3.4**).

The percentage of correctly detecting the second smallest red increment (1.26 cd/ m^2) may also contribute to CVD discrimination. To test this idea, we computed a weighted linear compound classifier combining the two smallest increments (**Table 3.4**). This blended more toward the upper left corner of the plot ROC curve, yielding an area under the curve of 0.906 when the weighting factor was $w = 0.7$. (red curve, **Figure 3.6**). Thus, using two red increments improves the red test's ability to classify CVDs, when compared to the HRR test.

Table 3.4. Sensitivity and Specificity Values for Red Test ROC Analysis

Smallest Increment (0.75 cd/m^2)

Portion Correct	Sensitivity	Specificity	Youden Index	Distance Index
0.1	0.0	1.0	0.0	1.0
0.15	0.286	0.971	0.257	0.715
0.2	0.286	0.971	0.257	0.715
0.25	0.643	0.886	0.529	0.375
0.3	0.786	0.771	0.557	0.313
0.35	0.786	0.771	0.557	0.313
0.4	0.786	0.771	0.557	0.313
0.45	0.786	0.743	0.529	0.335
0.5	0.786	0.743	0.529	0.335
0.55	0.857	0.714	0.571	0.319
0.6	0.857	0.714	0.571	0.319
0.65	0.929	0.629	0.557	0.378
0.7	0.929	0.629	0.557	0.378
0.75	0.929	0.486	0.414	0.519
0.8	0.929	0.486	0.414	0.519
0.85	1.0	0.371	0.371	0.629
0.9	1.0	0.371	0.371	0.629
0.95	1.0	0.257	0.257	0.743

Two Smallest Increments (0.75 ⨭ 1.26 cd/m^2)

Portion Correct	Sensitivity	Specificity	Youden Index	Distance Index
0.1	0.0	1.0	0.0	1.0
0.15	0.0	1.0	0.0	1.0
0.2	0.143	1.0	0.143	0.857
0.25	0.286	1.0	0.286	0.714
0.3	0.357	0.971	0.329	0.644
0.35	0.357	0.971	0.329	0.644
0.4	0.429	0.971	0.4	0.572
0.45	0.643	0.914	0.557	0.367
0.5	0.786	0.886	0.671	0.243
0.55	0.857	0.771	0.629	0.270
0.6	0.857	0.771	0.629	0.270
0.65	0.857	0.743	0.6	0.294
0.7	0.929	0.743	0.671	0.267
0.75	0.929	0.657	0.586	0.350
0.8	0.929	0.6	0.529	0.406
0.85	1.0	0.457	0.457	0.543
0.9	1.0	0.371	0.371	0.629
0.95	1.0	0.257	0.257	0.743

Optimal portion correct was analyzed with two quality indices computed for the tabluated data:
Youden Index = sensitivity + specificity – 1, higher is better (Youden 1950).
Distance Index = $\sqrt{((1 - \text{sensitivity})^2 + (1 - \text{specificity})^2)}$, the distance from the curve to the upper left corner of the plot, lower is better (Perkins & Schisterman 2006).

CHAPTER 4

Discussion

4.1 HRR: Normals vs. CVDs

The HRR has been a widely used assessment to test color vision and used in comparison with various other test for many years. The original assessment was developed in 1955 and revised in 2002 with research indicating that with moderate and severe defects the HRR is very accurate in categorizing subjects as protan or deutan (Bailey, Neitz, Tait, and Neitz, 2004). More recently, using ROC curve analysis that indicates the degree of discrimination to compare tests, it was found that the HRR test was the most successful with discriminative accuracy of color vision testing in cone disorders between controls with normal vision and subjects with abnormal color vision (Thiadens, et al, 2013). Discriminative accuracy defined as the area under ROC curve was 1.0 for the HRR plates along the protan/deutan axes compared to Ishihara plates (0.985), the Lanthony D-15 panel (0.890),and the Farnsworth D-15 panel (0.756). For the tritan axis, discrimination was 0.851 for the HRR plates. Within the Thiadens et al. 2013 ~~is~~ study, subjects with poor vision or legal blindness performed even less well across these 4 tests. Although the HRR plates are commonly the most accurate in color deficiency evaluation, the studies cited above note that the HRR is not error free.

In our current study, the HRR was our standard for determining the classification of normal or color vision deficient subjects, then used in comparison to the color

board/crayon, paint chips, and red test to gauge how sensitive these tests were and whether they might be used in the future for mass screenings, especially for children. For our normal population, there was a 37% error rate with the HRR, and all errors were made with plate 7 (subjects 15-49 were classified as normal by the HRR test despite these errors). A similar result was found by Yevseyenkov (2019) in his study tovalidate the 4th edition of the Hardy-Rand-Rittler where the error rate for normals was 14.2% with plate 7, an improvement over the 3rd edition which yielded 26.5% errors for plate 7 but still does not yield 100% sensitivity. York and Loop (2008) also found among their normal subjects that 16% failed plate 7 in the HRR test.

Birch (2010) determined that the vanishing figures for some of the plates had such small color differences that they can produce false positives in the normal populations which is why one missed figure is acceptable for a normal subject. This explanation is the potential cause for the prevalence of errors with plate 7 in the HRR test.

4.2 Crayon/Color Board: Normals vs. CVDs

Our findings with these two assessments were that identification errors were prevalent for both normals and CVD subjects, suggesting that physical properties of the samples made color discrimination difficult and failed to detect CVDs. Difficulties could arise from desaturated samples, especially with the color board test, or from imperfectly achromatic composition of the crayons themselves. For the normals, while there were only 2 errors (6%) on the crayon test, 8 errors (23%) were made on the color board test; all of these errors were in misidentifying achromatic colors for chromatic. Crayons are made with paraffin wax which when used on white paper causes deposition

to be inconsistent and allows the white background to show through to varying degrees, thereby making the test sample appear desaturated. With the crayons, the mixing of wax and the powdered pigments (Crayola Crayon, email communication, April 28, 2021) may not be perfect for achromatic crayons, and darker colors can be difficult to distinguish because of low luminance. Both the crayon and color board tests were compared to the HRR and found to be of low utility. These basic findings are consistent with a study on children (712 subjects) who were tested with the use of the Ishihara plates and crayon-colored squares on a sheet (Martins, et.al, 2001). The children were asked to color in the empty square with the correct color matching the one above it. The results of this test were that it also had low sensitivity of 38.5% (15.1-67.7 confidence intervals) and was deemed inappropriate for color vision screening.

For the CVDs, results are clear that our crayon test does not discriminate well between CVDs and normals. For the small number of subjects in the medium deficit range (n = 2), it is possible that by chance these subjects made no mistakes given the low error rate among mild CVDs in the HRR test. Subjects 13 and 14 (strong CVD) had performance that was consistent between the HRR and crayon tests, as they confused several achromatic colors with chromatic colors. With the color board test, there were 4 CVDs (across mild, medium, and strong) who made errors (29%) on the classification of colors, whereas there were 7 normals (20%) with classification errors for this test, again suggesting that the desaturation of colors from deposition on a white background can potentially make it difficult for anyone to classify. Thiadens et al. (2013) developed a comparative study assessing the accuracy of the HRR, the Ishihara, the Farnsworth D-15, and the Lanthony D-15 among the same group of adults. The Lanthony D-15 is a

desaturated color hue test that resembles the Farnsworth D-15 in application. Their findings for the Lanthony D-15, were that errors were more prevalent (up to 84% of their subjects had two or more crossings) compared to the Farnsworth D-15 (54% with two or more crossings), and that 44% of control subjects scored as abnormal. This suggests that desaturated tests can produce a high number of false positives, which is consistent with our findings of high error rates for the color board stimuli which have a desaturated appearance.

In counterpoint to the idea that normals can fail the currently available color discrimination test, it is also worth noting that dichromats can have relatively good performance with activities such as color naming, possibly because of individual differences with their cone mosaic (Cole.et.al, 2006). Cicerone (1990) studied the color appearance and cone mosaic in both trichomats and dichromats and determined that the density and number of cones in the fovea of a dichromat should be equal to the cones in the fovea of a trichomat. There is research indicating that some dichromats use rods to compensate for the lack of the third cone (Green, 2004). Scheibner and Boynton (1968) also concluded that the majority of dichromats preserve a weak functioning red-green chromatic system. The ability of dichromats to have good color classification under circumstances when samples reflect light that is predominantly of one wavelength may arise from their ability to use "brightness" cues to discriminate wavelengths (see Figure 5.1). With wavelengths away from their respective neutral points, dichromats might perceive brightness differences for lights that are otherwise matched in luminance for trichromats.

4.3 Paint Chips: Normals vs. CVDs

The normal subjects had no difficulty in identifying colors using the paint chips board as no errors were produced, while a few mistakes were made by the CVDs indicating that the paint chips are not great classifiers for color vision. Research conducted by Cole, Lian, Sharpe, and Lakkis (2006) conducted testing similarities in that they had 100 CVD participants and 20 normals who were asked to name 10 colors (orange, red, brown, yellow, green, black, blue, purple, white, and gray). These colors were presented to the subjects as dots and lines and with 3 different sizes under controlled illumination. Subjects were given the 10 color names of the stimuli and told to only use those names. They were given 2 seconds to name each of the colors, but the time limit was rarely needed. Diagnosis of CVD was done using the Nagel anomaloscope, Farnsworth D-15, the Medmont C100, and the Ishihara. Color classification errors were measured for both those who passed the Farnsworth D-15 and those who did not. Our interest here is on those who did not pass the Farnsworth D-15. Their data showed that certain chromatic samples had high error rates while achromatic ones had low error rates, as follows: red 15%, orange 13%, brown 32%, yellow 0%, green 12%, blue 3%, purple 4%, grey 3%, black 1%, and white 0.05%. They concluded that mild deuteranormals will make few errors if orange, brown, and purple were eliminated from the 10 color sets, as those were most commonly confused with another color (see their Fig. 8). These data are consistent with our findings that subjects in the mild HRR category sometimes misclassified chromatic paint chips. It is also worth noting that error rates in this task decrease as sample dimension increases (Cole, Lian,

Sharpe, and Lakkis, 2006), suggesting that stimulation of more retinal area helps avoid the potential problem of local sampling of the cone array.

Montag and Boynton (1987) examined surface color perception or color naming and found that several dichromats could categorize colors by name with accuracy close to that of normal subjects due to the contribution of rod signals and lightness cues. We found that dichromats could categorize some colors with complete accuracy (**Table 3.2**) when using the three research color tests (color board, crayons, paint chips). However, Montag (1993) created a surface color naming experiment and found that "an anomalous third cone pigment is responsible for the color categorization in three dimensions." He concluded that dichromats' color categorization may have contributed to greater temporal and spatial summation within receptors containing anomalous pigment.

Given our experiment using paint chips to screen for color vision deficiency we found that paint formulas would have to be developed and standardized to specifically align with the confusion lines for CVDs on the CIE chromaticity chart.

4.4 Red Test: Normal vs. CVDs

By definition, the spectral sensitivity function is the efficacy of detection of light as a function of wavelength. The red test was designed to measure sensitivity by measuring increment thresholds of red light on a gray background.

Suggestions from prior research imply that the use of a neutral background and spatially restricted increments could potentially help decide if a luminance or chromatic mechanism satisfy the detection threshold spectral increments in CVDs (Loop, Shows,

Mangel, and Kuyk, 2003) (Diaconu, Sullivan, Bouchard, and Vcea, 2010). Kaiser (2005) stated that there are 8 typical ways to measure spectral sensitivity in humans: (a) forced choice preferential looking, (b) electrophysiology, (c) minimally distinct border, (d) increment threshold, (e) absolute threshold, (f) critical flicker frequency, (g) flicker photometry, and (h) brightness matching. When choosing the best psychophysical measure for this research project, we considered each of these eight measures to assure the chosen technique would render the ability to produce quick mass screening. The forced choice preferential looking is typically used to check visual development and color vision in infants using a 2 forced-choice psychophysical technique (Brown and Lindsey, 2013), a technique not needed for school-age children. The electrophysiology for visual assessment uses a number of elaborate techniques for measuring cell function along the visual pathway (de Monasterio, 1984), all requiring a considerable amount of equipment that would only be suitable for one person at a time. Minimally distinct border is a method that evaluates relative radiance of two different color fields until the edge between fields is minimally detectable (Kaiser, Lee, Martin, and Valberg, 1990), a potentially demanding task for children. Absolute threshold is the smallest stimulus level detectable (Koenig and Hofer, 2011), and is known to be a challenging task. The critical flicker frequency is the lowest frequency at which a flickering light appears steady (Brenton, Thompson, and Maxner, 1989), something that could again only screen one person at a time. The same issue holds for flicker photometry, method that measures spectral sensitivity by using two different colored lights alternated and intensities adjusted by the subject (Lee, Martin, and Valberg, 1988), and brightness matching where intensity adjustments of two stimuli of different colors are matched (Fotios and Cheal,

2010). Thus, all of these established techniques would require specialized equipment or be time intensive for screening purposes as individual testing is necessary.

We chose to use the increment threshold measures with five-alternative forced-choice method using red increments with a commonly available classroom projection system. The normal subjects achieved relatively good scores for this task and exhibited suprathreshold performance even on the smallest increment. The CVDs' performance was lower on average, and they resorted to guessing on the lowest increment due to stimulus undetectability. In **Figure 3.4,** shows a difference in performance between the two groups across all increment levels. The data show that performance for CVDs across nearly all increments was worse than for normals, and the strongest CVDs performed worse than all other subjects with all increment levels. This indicates that the red test was consistent with the HRR classification, and even though there were limitations (to be discussed below), we believe that most CVD subjects can be detected with the red test, and that subjects with the most severe CVD, compared to the normal population, were not the source for the difference in the population CVD performance levels.

4.5 Limitations and Methodological Improvements

The shortcomings of this project would first be the sample size. The number of subjects in the medium and strong HRR groups were too small to allow any definitive conclusions to be drawn. For the purpose of replication, some potential issues with the color board/crayon and paint chips boards are that they can potentially become soiled and the oils from the skin will change the colors over time. This is also the case with the HRR plates. Another improvement would be to choose crayons/paint chips that would

fall along the color confusion lines for protanopes and deuteranopes in the CIE diagram. Other potential weaknesses are the relying on color formulas that may change for commercial color samples and the known fading of samples from extended light exposure. For instance, according to Hyon, Lee, and Wee (2005), the colorimetric values significantly differ due to plate aging, soiling, and fading. They made spectrophotometric measurements of pseudoisochromatic plates with various publication dates and found significant changes in the color values causing a CIE directional shift with the aging process.

In the research conducted by Cole, Lian, Sharpe, and Lakkis (2006) we observed that they did not pick many colors on a single line using the CIE diagram. They suggested that by using 7 specific colors, the error rate for mild deuteranormals could become quite small (average around 0.3%), but this would not be the case with other color deficiency classes. With future studies using the paint chips, specific paint formulas would have to be created to be on the color confusion lines. Of course, this would eliminate the utility of using commercially available color samples for color deficiency testing.

When performing the red test, it is wise to consider the type of projection system being used such as LCD (liquid-crystal display) vs. DLP (digital light processing). LCD delivers better color saturation in commercial projectors, but LCD has a long-term image degradation, meaning that the LCD panels and polarizers will degrade over time (Powell, 2005). In DLP projectors, light from a bulb of some type is modulated via a micromirror device which then passes through a color filter wheel. These components may become faulty or spectrally shifted over time. Specifications such as lumen output, increment

threshold settings depending on the psychophysical measure, and the calibration of the light meter would be needed to ensure increments as low as 1 cd/m^2 could be delivered. Also, it would be useful to consider the type of lighting within the environment for all the tests, such as florescent vs. natural light, since ambient lighting can change the appearance of color (Wandell, 1987). For example, in a room with only fluorescent lighting, the small number of narrow spectral peaks in this source may shift the color appearance of a stimulus relative to a room illuminated with natural light (Green, 2004).

4.6 Future Studies

As for future studies, it would be of interest to see how children would perform on these same tests in a mass screen setting. In the study by Nietz and Nietz (2001), they expounded on the fact that color coding is used daily to communicate and teach basic curriculum. With the extensive use of color in the early education of children, there would be obvious benefits for screening children at an early stage. Accommodations should be made in school for their educational success and ultimately future career choices and to eliminate over-identification of learning disabilities.

From the results of our study, we determined that the red test could discriminate about as well as the HRR. Our methods using red increments with adults could be used as a test for children, extending the results of York and Loop (2008) who found that red increment thresholds were higher in CVDs. This method, capitalizing on insensitivity to red, could even be implemented with preferential looking techniques in infants (Vital-Durand and Cottard, 1985).

Currently, the test is a slide presentation with 50 trials, measuring detection across 5 increment levels. When looking at the results for both CVDs and normals, we would likely implement only 3 increments as we found that some of the increment levels were redundant and could be eliminated for future testing. It would be best if at least one increment would be above threshold for all subjects, and one would be below threshold for normals. The number of trials per condition could be reduced to 5. This brings the test to a total of 15 trials, making it more manageable when testing children. To redesign the red test to be appropriate for children, the forced choice method with gaussian noise on a gray background would still be used but shapes (squares or triangles) could replace the letters A-E to accommodate smaller children. We could also base the answer sheet ona yes or no response as this would allow children to be screened by classrooms instead of individually.

Alternatively, we could also use printed cards similar to the paper and pencil test developed by Nietz and Nietz (2000), where children traced over colored symbols with crayons or pencils directly on test papers. These papers were comparable to both the Ishihara, the HRR, and the APT-5 Color Vision Tester patterns. They determined that this particular method was quick and reliable to test groups of children for color vision. Teachers administered the test to 20-30 children at the same time and children who madeone or more errors were re-tested with the Nietz paper and pencil test. Children who made errors on the Nietz test for a second time were tested with one or more of the conventional color deficiency tests. With their approach, 18% of the children made one or more errors on the first administration, once re-tested they found that all the children who were classified as CVD by the paper and pencil test also failed one or more of the conventional tests.

The testing time per child was less than 2 minutes, which also indicates quality management of time when testing children. In adapting our red test for children, we aim to shorten the time spent assessing children to 2-3 minutes per child or group of children.

Further research for color vision screening with materials that are easily available and allow for mass testing, especially with children, should begin with the basic pre-school age of 4-5 years, with a re-test at around 8-9 years since color discrimination improves with age to a certain point in life. According to Knoblauch et al. (2001), chromatic thresholds decreased with the doubling of age until age 20, at which time the chromatic thresholds begin to increase. However, testing would potentially provide the information needed to make informed decisions concerning education, technology, careers and environments accessible for success starting at an early age.

References

Almustanyir A. Review of the Color Vision Tests Currently in Use. J Ophthalmology & Vis. Sci. 2020; 5(1): 1036

Anthony J and Spalding B (2004) Confessions of a colour blind physician, Clinical and Experimental Optometry, 87:4-5, 344-349

Bailey, J., Neitz, J., Tait, D, and Neitz, M. (2002). Updating the HRRColorVisionTest. 3:00 -3:15 Optometry and Vision Science, 79(12), 282.

Bailey JE, Neitz M, Tait D, and Neitz J. Evaluation of an updated HRR color vision test. 2004 May-Jun;21(3):431-6. PMID: 15518225

Birch J. Identification of red-green colour deficiency: sensitivity of the Ishihara and AmericanOptical Company (Hard, Rand and Rittler) pseudo-isochromatic plates to identify slight anomalous trichromatism. Ophthalmic Physiol Opt. 2010 Sep;30(5):667-71. PMID: 20883353.

Blais, B. (2010) A Further Word on Color Vision and its Relevance in the Workplace. Color Vision Deficiency. Highley.

Brenton R.S., Thompson H.S., Maxner C. (1989) Critical Flicker Frequency: A New Look at an Old Test. In: Wall M., Sadun A.A. (eds) New Methods of Sensory Visual Testing. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-8835-7_3

Brown AM, Lindsey DT, Guckes KM. Color names, color categories, and color-cued visual 2. search: sometimes, color perception is not categorical. J Vis. 2011 Oct . 6;11(12):10.1167/11.12.2 doi: 10.1167/11.12.2. PMID: 21980188; PMCID: PMC3632453

Chapanis, A. (1944). Spectral saturation and its relation to color-visiondefects. Journal of Experimental Psychology, 34(1), 24.

Chaparro, A., Stromeyer, C. F., Huang, E. P., Kronauer, R. E., & Eskew, R. T. (1993). Color is what the eye sees best. Nature .

Cicerone, C. Color Appearance and the Cone Mosaic in Trichromacy and Dichromacy. E-Scholarship University of California, 1990.

Cole, B. L. (2004). The handicap of abnormal color vision. Clinical and Experimental. Optometry, 87(4-5), 258-275.

Cole BL, Lian KY, Sharpe K, Lakkis C. Categorical color naming of surface color codes bypeople with abnormal color vision. Optom Vis Sci. 2006 Dec;83(12):879-86. doi: 10.1097/01.opx.0000249974.22205.2a. PMID: 17164680.

Cole, B. L., & Orenstein, J. M. (2003). Does the Farnsworth D15 test predict the ability. to name colors? Clinical and Experimental Optometry, 86(4), 221-229.

Coren, S., & Hakstian, A. R. (1988). Color vision screening without the use of Technical equipment: Scale development and cross-validation. Perception & Psychophysics, 43(2), 115- 120.

Cranwell, M. B., Pearce, B., Loveridge, C., & Hurlbert, A. C. (2015). Performance on the Farnsworth-Munsell 100-Hue Test Is Significantly Related to Nonverbal IQFM100 Test Significantly Related to NVIQ. Investigative ophthalmology & visual science, 56(5), 3171-3178.

Crayola Crayon (email communication) April 28, 2021

Curcio, C.A., Allen, K.A., Sloan, K.R., Lerea, C.L., Hurley, J.B., Klock, I.B. and Milam, A.H.(1991), Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. J. Comp. Neurol., 312: 610-624.

Dain SJ, Atchison DA, Hovis JK. Limitations and Precautions in the Use of the Farnsworth-Munsell Dichotomous D-15 Test. Optom Vis Sci. 2019 Sep;96(9):695-705. doi: 10.1097/OPX.0000000000001420. PMID: 31479025.

Dain, S. J., & King-Smith, P. E. (1981). Visual thresholds in dichromats and normals; the importance of post-receptoral processes. Vision research, 21(4),573-580.

Deeb, S.S. (2004), Molecular genetics of colour vision deficiencies. Clinical and ExperimentalOptometry, 87: 224-229.

Deeb, S. S. (2005). The molecular basis of variation in human color vision. Clinical genetics, 67(5), 369-377.

Deeb SS, Motulsky AG. (2005) Red-Green Color Vision Defects. Seattle (WA): Gene Reviews, University of Washington, Seattle; 1993-2017

de Monasterio F.M. (1984) Electrophysiology of Color Vision. In: Verriest G. (eds) Colour Vision Deficiencies VII. Documenta Ophthalmologica Proceedings Series, vol 39. Springer, Dordrecht.

Diaconu V, Sullivan D, Bouchard JF, Vucea V. Discriminating colors through a red filter byprotanopes and colour normals. Ophthalmic Physiol Opt. 2010 Jan;30(1):66-75. doi: 10.1111/j.1475-1313.2009.00695.x. PMID: 20444111.

Fotios S, Cheal C. Brightness matching with visual fields of different types. Lighting Research& Technology. 2011;43(1):73-85. doi:10.1177/1477153510369478

Gallo, P. G., Panza, M., & Viviani, F. (1998). Congenital dyschromatopsia and school achievement. Perceptual and motor skills, 86(2), 563-569.

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. Annual review of neuroscience, 26(1), 181-206.

Green, Marc (2018) Perception and Human Factors. *Human Factors*. *https://www.visualexpert.com/why.html*

Hardy, L. H., Rand, G., & Rittler, M. C. (1954). H–R–R Polychromatic Plates. *JOSA*, *44*(7),509-523.

Hecht, S., & Shlaer, S. (1935). The Color Vision of Dichromats I. Wavelength Discrimination, Brightness Distribution, And Color Mixture. The Journal of general physiology, 20(1), 57-82.

Hosmer, D. Jr., Lemeshow, S., and Sturdivant, R. (2013) Applied Logistic Regression, 3rd Edition, pp177

Hsia, Y., & Graham, C. H. (1957). Spectral luminosity curves for protanopic, deuteranopic, andnormal subjects. Proceedings of the National Academy of Sciences, 43(11), 1011-1019.

Hyon, J. Y., Lee, J. H., & Wee, W. R. (2005). Shift of colorimetric values in Ishihara pseudoisochromatic plates with plate aging. Korean Journal of Ophthalmology, 19(2), 145-148.

Jägle, H., Pirzer, M., & Sharpe, L. T. (2005). The Nagel anomaloscope: its calibrationand recommendations for diagnosis and research. Graefe's Archive for Clinical and Experimental Ophthalmology, 243(1), 26-32.

Jurasevska, K., Ozolinsh, M., Fomins, S., Gutmane, A., Zutere, B., Pausus, A., and Karitans, V. (2014). Color-discrimination threshold determination using pseudoisochromatic test plates. Frontiers in psychology, 5.

Kaiser, P. K., Lee, B. B., Martin, P. R., & Valberg, A. (1990). The physiological basis of theminimally distinct border demonstrated in the ganglion cells of the macaque retina. The Journal of Physiology, 422(1), 153-183.

Kaiser, P. K. (2009). *The joy of visual perception*. Merlot Multimedia Educational Resource for Learning and Online Teaching, York University.

King-Smith, P. E., & Carden, D. (1976). Luminance and opponent-color contributions visual detection and adaptation and to temporal and spatial integration. JOSA, 66(7), 709-717.

Kintz, R. T. (1983). A portable, solid-state anomaloscope. Behavior Research Methods& Instrumentation, 15(6), 587-590.

Koenig, D., & Hofer, H. (2011). The absolute threshold of cone vision. *Journal of vision*, *11*(1),21-21.

Knoblauch, K., Vital-Durand, F., & Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. Vision research, 41(1), 23-36.

Lee, B. B., Martin, P. R., & Valberg, A. (1988). The physiological basis of heterochromatic flicker photometry demonstrated in the ganglion cells of the macaque retina. The Journal ofphysiology, 404(1), 323-347.

Ling, B. Y., & Dain, S. J. (2008). Color vision in children and the Lanthony New Color Test. Visual neuroscience, 25(3), 441.

Long, J. A., Honson, V., Katalinic, P., & Dain, S. J. (2015). Re: Is screening forcongenital color vision deficiency in school students worthwhile? A review Clinical and Experimental Optometry, 98(2), 192-192.

Loop, M. S., Shows, J. F., Mangel, S. C., & Kuyk, T. K. (2003). Colour thresholds in dichromatsand normals. Vision research, 43(9), 983-992.

Montag, E. D. (1994). Surface color naming in dichromats. *Vision Research*, *34*(16), 2137-2151.

Montag, E. D., & Boynton, R. M. (1987). Rod influence in dichromatic surface color perception. Vision Research, 27(12), 2153-2162.

Neitz, M., & Neitz, J. (2000). Molecular genetics of color vision and color vision defects. Archives of Ophthalmology, 118(5), 691-700.

Neitz, M., & Neitz, J. (2001). A new mass screening test for color-vision deficiencies in children. Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, ColourSociety of Australia, Centre Français de la Couleur, 26(S1), S239

Oliveira, M. M. (2014). Uncertainty visualization and color vision deficiency. In ScientificPasmanter, N., & Munakomi, S. (2020). Physiology, color perception. StatPearls Visualization (pp. 29-33). Springer London.

Perkins N. J., Schisterman E. F. The inconsistency of "optimal" cut-points using two ROC basedcriteria. American Journal of Epidemiology. 2006;163(7):670–67

Pitt, F. H. G. (1935). Characteristics of dichromatic vision with an appendix on anomalous Trichromatic vision, Med. Res. Council. Spec. Rep. Ser, (200).

Powell, E. (2005) The Great Technology War: LCD vs. DLP. *Projector Central. https://www.projectorcentral.com/lcd_dlp_update.htm*

Pramanik, T., Sherpa, M. T., & Shrestha, R. (2010). Color vision deficiency among medical students: an unnoticed problem. Nepal Med Coll J, 12(2), 81-83.

Pridmore, R. W. (2014). Orthogonal relations and color constancy in dichromatic colorblindness. PLoS One, 9(9), e107035.

Ramachandran, N., Wilson, G. A., & Wilson, N. (2014). Is screening for congenitalcolor vision deficiency in school students worthwhile? A review a. Clinical and Experimental Optometry, 97(6), 499-

Scheibner, H. M., & Boynton, R. M. (1968). Residual red-green discrimination in dichromats. JOSA, 58(8), 1151-1158

Schroeder, M. M., Harrison, K. R., Jaeckel, E. R., Berger, H. N., Zhao, X., Flannery, M. P., & Wong, K. Y. (2018). The roles of rods, cones, and melanopsin in photoresponses of M4 intrinsically photosensitive retinal ganglion cells (ipRGCs) and optokinetic visual behavior. Frontiers in cellular neuroscience, 12, 203.

Schwartz, B.L. (2015) *Color Perception.* Sensation and Perception. SAGE Publications.

Schwartz, B. L., & Krantz, J. H. (2015). Sensation and Perception. SAGE Publications.

Schwartz, S.H. (1994) Spectral sensitivity of dichromats: Role of postreceptoral Processes. Vision Research, 34 (1994), pp. 2983–2990

Schwartz, S. H. (2010). Visual perception: A clinical orientation. New York: McGraw-Hill Medical Optometry, 87(4-5), 258-275.

Sincich, L. C., Sabesan, R., Tuten, W. S., Roorda, A., & Harmening, W. M. (2016). Functional imaging of cone photoreceptors. In Human color vision (pp. 71-104). Springer, Cham

Spalding, J. A. B. (1997). Doctors with inherited colour vision deficiency: their difficulties inclinical work. In Colour vision deficiencies XIII (pp. 483-489). Springer, Dordrecht.

Spalding, J. A. (1999). Color vision deficiency in the medical profession. Br J Gen Pract, 49(443), 469-475.

Spalding, J. A. B. (2004). Confessions of a colour blind physician. Clinical and ExperimentalOptometry, 87(4-5), 344-349.

Sperling, H. G., & Harwerth, R. S. (1971). Red-green cone interactions in the increment threshold spectral sensitivity of primates. Science, 172(3979), 180-184

Steinkuller, P. G. (2010). Legal vision requirements for drivers in the United States. *AMA* Journal of Ethics, 938-940.

Suckow, M., & Watson, T. (2002). A State-By-State Analysis of Preschool Vision ScreeningPolicy in The United States.: Poster# 44. Optometry and Vision Science, 79(12), 226.

Stockman, A., & Sharpe, L. T. (2000). The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype.Vision research, 40(13), 1711-1737.

The National Center for Children's Vision and Eye Health (2020). Health Resources and Services Administration. https://www.hrsa.gov/library/national-center-childrens-vision-eye-health-0.

Thiadens AA, Hoyng CB, Polling JR, Bernaerts-Biskop R, van den Born LI, Klaver CC. Accuracy of four commonly used color vision tests in the identification of cone disorders. Ophthalmic Epidemiol. 2013 Apr;20(2):114-21. doi:10.3109/09286586.2012.759596. PMID:23510316.

Verriest, G., & Uvijls, A. (1977). Central and peripheral increment thresholds for white

Vital-Durand F, Cottard M. La technique du regard préférentiel: matériel et procédure de bébé vision [The preferential looking technic: material and procedure for testing baby

Wandell, B.A. "The Synthesis and Analysis of Color Images," in *IEEE Transactions or*

Pattern Analysis and Machine Intelligence, vol. PAMI-9, no. 1, pp. 2-13, Jan. 1987, doi: 10.1109/TPAMI.1987.4767868.

Watson AB, Barlow HB, Robson JG. What does the eye see best? Nature. 1983 Mar 31-Apr6;302(5907):419-22. doi: 10.1038/302419a0. PMID: 6835375.

Whillans MG, Allen MJ. Color defective drivers and safety. Optom Vis Sci. 1992 Jun;69(6):463-6. doi: 10.1097/00006324-199206000-00009. PMID: 1641229.

Yamaguchi, T., Motulsky, A. G., & Deeb, S. S. (1997). Visual pigment gene structure and expression in human retinae. Human Molecular Genetics, 6(7), 981-990.

Yevseyenkov, V. (2019) Validation of Fourth Edition Hardy-Rand Rittler PseudoisochromaticTest. *American Academy of Optometry.* Orlando, Fl.

Youden, W.J. (1950*). "Index for rating diagnostic tests". Cancer. 3: 32–35.*

York, Y. C., & Loop, M. S. (2008). Red light increment threshold as a measure of deficient color vision. Optometry and Vision Science, 85(2), 106-111.

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.*

APPENDIX A

CIE CHROMATICITY COORDINATES

APPENDIX A

CIE CHROMATICITY COORDINATES

CIE chromaticity coordinates were measured from the samples as 2° observer values with a Photoresearch PS-670 spectrophotometer in a classroom with fluorescent lights. A white standard measured in the room had CIE coordinates of x = 0.411, y = 0.408, color temperature of 3510 K (slightly "yellow"), and luminance of 96 cd/m^2.

Paint Chips

	CIE x	CIE y
Pink Taffy 2075	0.439	0.366
Blue Orchid 2069	0.399	0.384
Pilgrim Haze 2132	0.400	0.402
Rhododendron 2079	0.457	0.370
Bayberry 2080	0.458	0.377
Metallic Silver	0.405	0.405
Mystical Grape 2071	0.391	0.346
Fairy Tale Blue 2055	0.347	0.404
Caribbean Blue	0.319	0.388
Budgreen 2033	0.385	0.453
Black	0.403	0.404
Amethyst Cream 2071	0.403	0.367
Suprawhite P841	0.412	0.409
Baby Bay Blue 2056	0.351	0.410
Passion Pink	0.428	0.382
Gull Wing Gray 2134	0.404	0.406
Laguna Blue 2059	0.317	0.359

Crayons	**On Colorboard**		**Direct**	
	CIE x	CIE y	CIE x	CIE y
Orchid	0.416	0.384	0.449	0.347
Purple Mountain Majesty	0.402	0.391	0.384	0.380
Gray	0.409	0.406	0.395	0.402
Red Violet	0.435	0.379	0.456	0.388
Mauvelous	0.432	0.389	0.474	0.366
Timberwolf	0.409	0.408	0.399	0.405
Purple	0.408	0.373	0.415	0.397
Turquoise Blue	0.374	0.407	0.343	0.375
Bluegreen	0.357	0.404	0.362	0.384
Sea Green	0.392	0.433	0.368	0.472

Black	0.412	0.408	0.409	0.409
Wisteria	0.408	0.384	0.382	0.345
White	0.409	0.407	0.422	0.416
Robin Egg Blue	0.377	0.416	0.314	0.418
Lavender	0.421	0.392	0.452	0.376
Silver	0.408	0.407	0.406	0.408
Cerulean	0.356	0.395	0.368	0.367

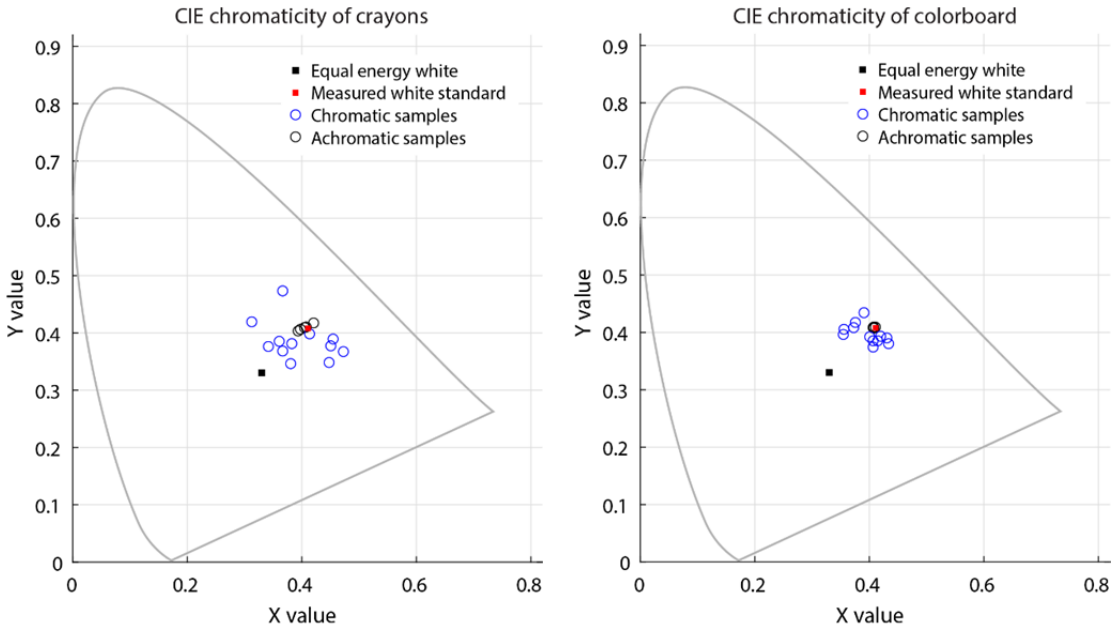


Figure A1. Crayon chromaticity values moved toward the "white" standard on the samples where crayon material was deposited on white paper. Mean luminance of the color board was also 3.3× brighter than the crayons (72 vs. 22 cd/m^2), mostly due to the transparency of the drawn crayon.

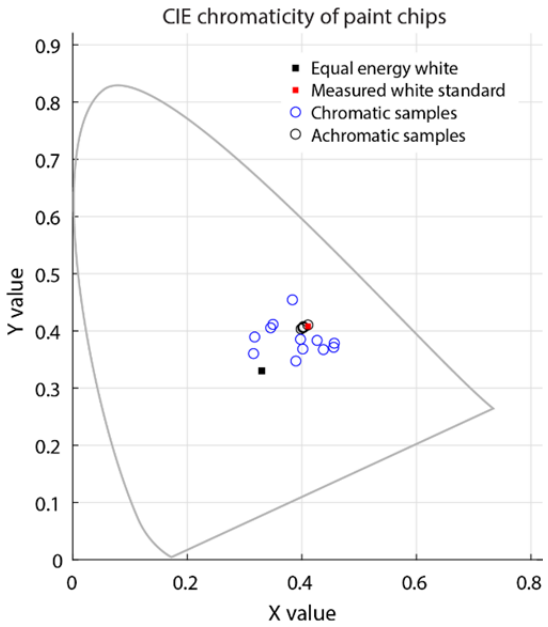


Figure A2. Paint chip chromaticity values. Mean luminance of the paint chips was 2.4× brighter than the crayons (51 vs. 22 cd/m^2).

APPENDIX B

UAB INSTITUTIONAL REVIEW BOARD APPROVAL

470 Administration Building701
20th Street South Birmingham, AL 35294-0104
205.934.3789 | Fax 205.934.1301 |
irb@uab.edu

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Office of the Institutional Review Board for Human Use

Approval Letter

To: Perez, Angel G

FROM: University of Alabama at Birmingham
Institutional Review Board Federal wide Assurance # FWA00005960
IORG Registration # IRB00000196 (IRB 01)
IORG Registration # IRB00000726 (IRB 02)
IORG Registration # IRB00012550 (IRB 03)

DATE: July 1, 2020

RE: IRB-300000548
Color Vision Testing with Readily Available Materials

The IRB reviewed and approved the Continuing Review submitted on Jun 15 2020 for the above referenced project. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services.

Type of Review: Expedited
Expedited Categories: 4,5,7,
Determination: Approved
Approval Date: Jun 16 2020
Expiration Date: Jun 15 2021

To access approved documents and/or the stamped consent/assent forms:

1. Open your protocol in IRAP.
2. On the Submissions page, open the submission corresponding to this approval letter. NOTE: The Determination for the submission will be "Approved."
3. In the list of documents, select and download the desired approved documents. The stamped consent/assent form(s) will be listed with a category of Consent/Assent Document (CF, AF, Info Sheet, Phone Script, etc.).