Temporal Color Perception

by

R. Chandler Krynen

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2021 by the
Graduate Supervisory Committee:

Michael McBeath, Chair
Donald Homa
Nathan Newman
Greg Stone

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Color perception has been widely studied and well modeled with respect to combining visible electromagnetic frequencies, yet new technology provides the means to better explore and test novel temporal frequency characteristics of color perception. Experiment 1 tests how reliably participants categorize static spectral rainbow colors, which can be a useful tool for efficiently identifying those with functional dichromacy, trichromacy, and tetrachromacy. The findings confirm that all individuals discern the four principal opponent process colors, red, yellow, green, and blue, with normal and potential tetrachromats seeing more distinct colors than color blind individuals. Experiment 2 tests the moving flicker fusion rate of the central electromagnetic frequencies within each color category found in Experiment 1 as a test of the *Where* system. It then compares this to the maximum temporal processing rate for discriminating direction of hue change with colors displayed serially as a test of the *What* system. The findings confirm respective processing thresholds of about 20 Hz for *Where* and 2-7 Hz for *What* processing systems. Experiment 3 tests conditions that optimize false colors based on the spinning Benham's Top illusion. Findings indicate the same four principal colors emerge as in Experiment 1, but at low saturation levels for trichromats that diminish further for dichromats. Taken together, the three experiments provide an overview of the common categorical boundaries and temporal processing limits of human color vision.

DEDICATION

To MP, for your motivation and inspiration.

ACKNOWLEDGMENTS

I would like to acknowledge the incredible and thoughtful assistance of my advisor, Dr. Michael McBeath, as I worked through this project. I would further like to thank my committee members for their insight, time and support. The technical assistance and wisdom of Steven Brown helped greatly with programming of the experiments and finally, the expertise offered by the SciHub Engineering and Physics group was deeply appreciated in creating much of the physical equipment. To my family and friends, thank you as well for your tremendous support throughout my time in graduate school.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER 1

RAINBOW STRIPES: CATEGORICAL PERCEPTION OF COLOR AS A TOOL FOR TESTING TETRACHROMACY

The potential for humans to have a fourth color cone type, known as tetrachromacy, has been demonstrated (Jameson, 2009; Jameson et al., 2020). Tetrachromatic individuals have the ability to discern a much wider range of colors than the average trichromat with normal vision, and has been likened to having "superhuman vision." However, estimates of prevalence of tetrachromacy vary widely between studies with estimates ranging from as little as 12% to as much as 50% in females (Jordan & Mollon, 1993; Backhaus et al., 1998; Jordan et al., 2010; Neitz et al., 1998). A major reason for the wide range of estimation has to do with tetrachromacy being difficult to test in an efficient manner; current displays and technology are optimized for trichromatic users, often employing a set of just three primary colors. Further, the ranges where tetrachromatic individuals have enhanced vision has also yet to be determined, with preliminary evidence pointing toward enhanced discrimination somewhere in the green-red range, 480-660nm. (Jordan & Mollon, 2019). These issues prompt the need for an efficient method that can both easily test many observers as well as give more precise locations of where trichromats and tetrachromats are likely see enhanced discrimination between colors.

Currently, a popular way to test for tetrachromacy is by using an expensive polymerase chain reaction (PCR) genetics test, previous work has demonstrated some humans, typically many females, do indeed possess a fourth color cone using this method

(Jameson et al., 2001). Unfortunately, the genetics test only confirms whether an individual has a fourth color cone type, but this does not necessarily mean that the fourth cone results in enhanced color vision discrimination. Individuals with a fourth color cone who do not behaviorally show tetrachromatic tendencies are known as being non-functional tetrachromats whereas those that have a fourth color cone type *and* show enhanced color discrimination are known as functional tetrachromats. Due to this, it would be useful to have a simple inexpensive method to test for functional tetrachromacy.

A relatively cost-effective test has been devised to test for functional tetrachromacy where a participant is given a standardized swatch of colors, the Munsell Color Checker (MCC), and then mixes paints to match the colors on the MCC as closely as possible (Bochko & Jameson, 2018). Once the participant completes the task, the painted swatch of colors is compared to the MCC using a spectrophotometer to determine how well the painted colors match the MCC. While a clever way for testing for functional tetrachromacy, this method has had limited acceptance for two reasons. One, the test is time-consuming and requires 2-2.5 hours for a participant to complete. Second, the experiment also requires a sense of artistic knowledge, knowing which paints to mix to produce a desired color.

Toward the goal of developing a more efficient way of testing for functional tetrachromacy, Jameson et al. (2001) displayed a full spectrum gradient of color to participants wherein perceived boundaries between colors were marked. For example, a participant marks down the edges of stripe colors for how many stripes they see (Figure 1.1). Functional tetrachromats typically indicated seeing more stripes than trichromats,

showing this to be a promising new way of testing for tetrachromacy. Some things to note about this method is that the stimulus itself is composed of expensive optical equipment to properly display the rainbow-like gradient to participants. Further, the stimulus was never verified to display uniform bands of color to participants and may have been artificially constrained to make some bands wider or narrower than others by adjustment of the various lenses used in the apparatus according to the experimenter's own visual perception, with a halogen lamp not producing a broadband spectrum of light of which other sources are capable. Finally, Jameson et al. note the best exemplar locations for trichromats and tetrachromats of where stripes were placed while being constrained to having to place seven stripes, but this amalgamation of data precludes the ability to elucidate where on average, new stripes are being placed between those participants that see just six stripes compared to those that see up to 16 stripes (no further analysis was done for a separate color demarcation task where participants were not constrained to six stripes). Determining where new stripes are being seen between participants can shed light on where tetrachromats have enhanced discernment between colors compared to their trichromatic counterparts. It becomes difficult to compare stripe delineations to best-exemplar locations in the Jameson et al. study as the instructions for each result were different: the former was a function of simple line demarcation as seen in Figure 1.1, whereas the latter used semantic color categories e.g. asking participants to label locations of red, orange, yellow, green, blue, violet, and purple so participants were constrained to just seven color choices. While this addressed the researcher's question of whether tetrachromats would place archetypal rainbow colors in different places

compared to trichromats--they did not--it does not answer the question of where tetrachromats are more likely to exhibit additional color discrimination.

The current study aims to extend the gradient demarcation idea to be more accessible, use a more ecologically-valid stimulus, and importantly identify the areas where functional tetrachromats will exhibit increased discernment between colors. If these parameters can be met, then one could feasibly more accurately determine prevalence of functional tetrachromacy as well as hone in more precisely on where tetrachromats show heightened sensitivity to color.

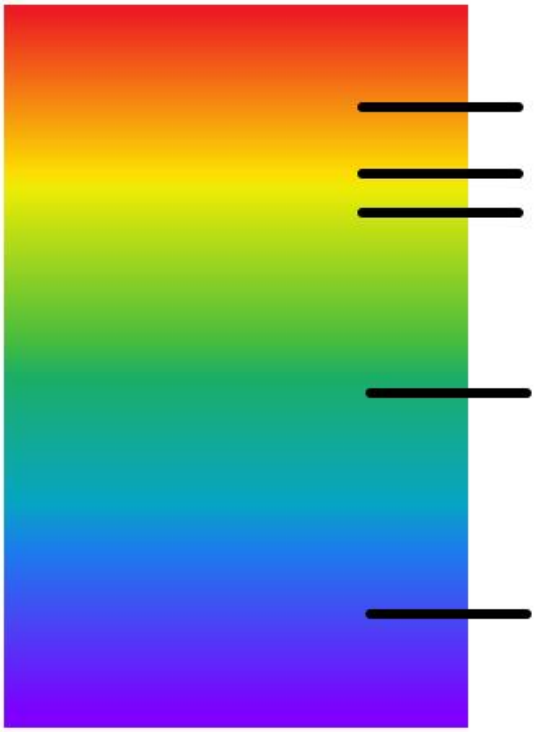

Figure 1.1. An example of how participants might delineate the edges of "stripes" of color in the Jameson et al. experiment (2001). Black bars represent a participant's final placement of lines indicating borders of the color stripes they see on a spectral rainbow pattern, similar to the one used in the current experiment.

One of the most prominent models used in color science is the Commission internationale de l'éclairage (International Commission on Illumination) or "CIE" 1931 color space first presented in Cambridge, United Kingdom (Jones, 1943), shown in Figure 1.2. This model geometrically transforms combinations of electromagnetic spectral frequencies into a three-dimensional mapping of perceived colors, created by two independent researchers, Guild (1932) and Wright (1929, 1930). In the original experiments, participants were given a reference color and instructed to match the reference color using three primary colors, a form of multidimensional scaling (MDS). However, to correct for differences of wattage of the primary lights as well as the human eye having inconsistent sensitivity to various wavelengths, participants first had to match the three primaries with a standardized 5000 kelvin National Physics Laboratory (NPL) Standard White Light. Once this procedure was performed, the experimenter had values that could then be multiplied by the subject's response to any test color resulting in the vector equation:

$$C \equiv \alpha R + \beta G + \gamma B$$

where α, β, and γ comprise the trichromatic coefficients. These coefficients specify what is now contemporarily called "chromaticity," which is the quality of color without regard to brightness. It was previously demonstrated that the human eye is most sensitive to light around 555nm with sensitivity diminishing at the shorter and longer wavelengths (CIE, 1926, Figure 1.3). To account for brightness, Guild created scaling factors (or in Guild's words, luminosity factors) using the equation:

$$L_R : L_G : L_B = 1 : 4.390 : 0.048$$

where each of the trichromatic coefficients is multiplied by these values and then summed to arrive at a luminance value that was associated with a perceived color. Ultimately, Guild and Wright succeeded in creating the framework for quantitatively describing visible colors, that has become the international color standard known as the CIE 1931 color space. The CIE color space comprises a multidimensional scaling (MDS) solution in which equal distances in in the space represent approximately equal perceptual differences. The Y axis represents relative luminance whereas the X axis is a function of all three cone responses.

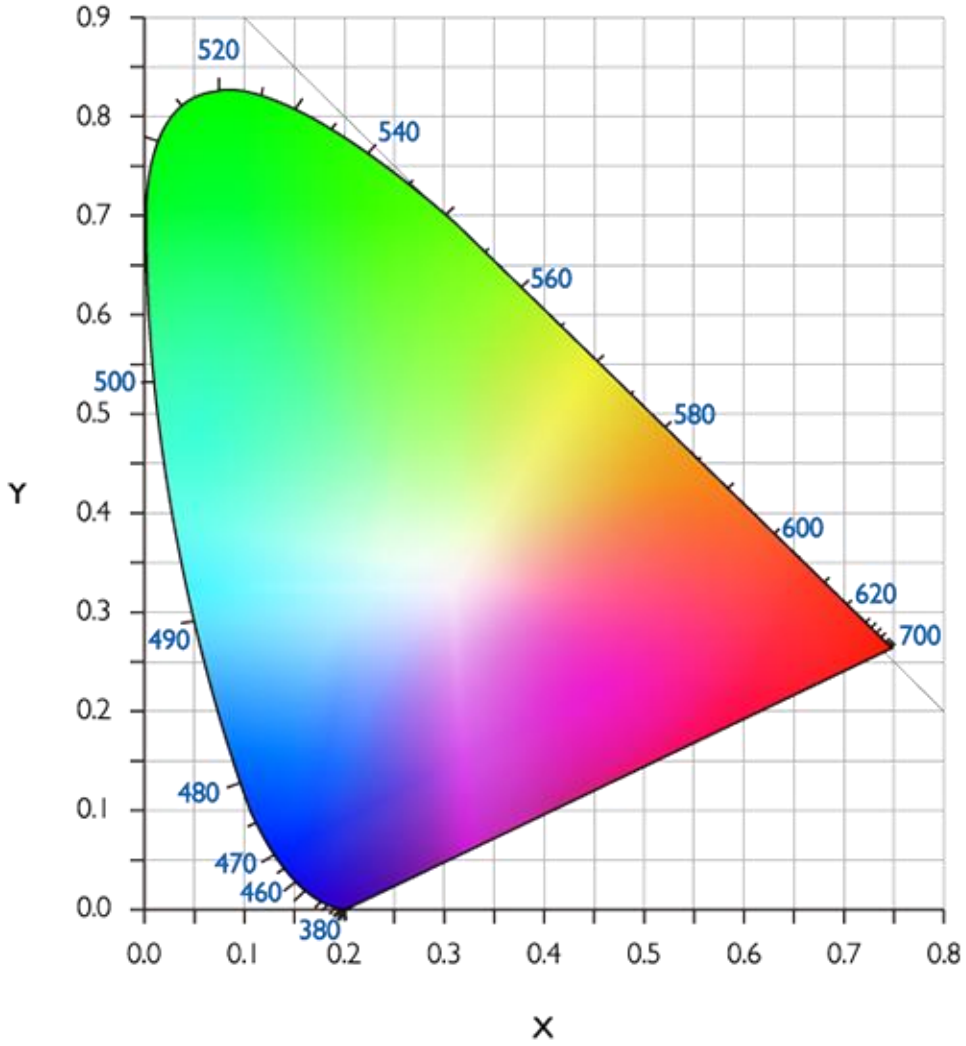

Figure 1.2. CIE 1931 color space. This space was created with matching three primary colors to monochromatic reference colors as a means of quantifying how the human eye perceives colors. The colors around the curved edge of the space are the pure wavelength spectral colors, and the colors within the space are multiple wavelength colors that can be approximately expressed as linear averages of their pure wavelength components. This is a classic multidimensional scaling (MDS) solution in which equal distances in CIE color space represent approximately equal perceptually-discriminable differences. The colors shown in this diagram are not actual CIE color space colors but rather are rough approximations for illustrative purposes.

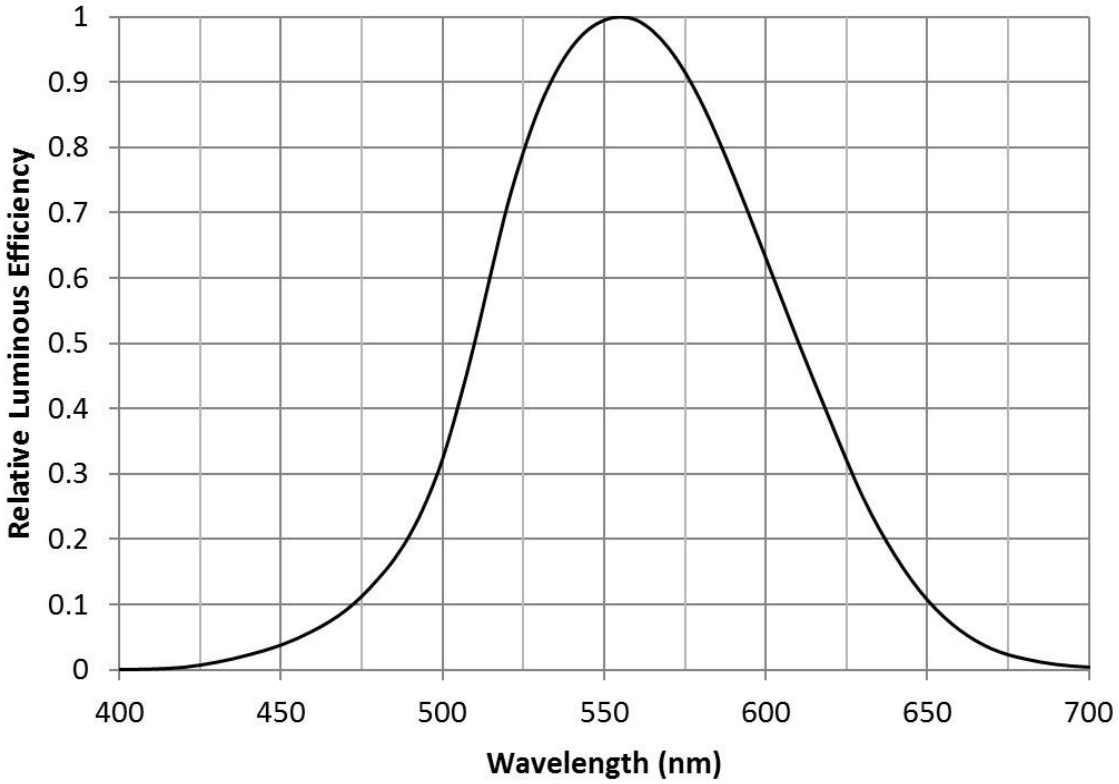

Figure 1.3. The human photopic luminous efficiency function, $V(\lambda)$ as defined by the CIE 1924 color commission. This function characterizes the sensitivity of the human eye at each wavelength where 555nm is the maximally sensitive wavelength. This function was derived using heterochromatic brightness matching. This is a simplified approximation of how an individual's retina would respond to various wavelengths of light.

The resultant CIE color space is relatively continuous, and appears to have no visible discontinuities or categorical stripes, unlike colloquial descriptions of a rainbow (Figure 1.4). In the study, I explore the idea of color categorization and test for prototypical regions of spectral colors that produce the perception of rainbow "stripes" in smooth gradients of pure wavelength color stimuli. Specifically, the wavelengths used in the current experiment will be pure wavelengths taken from the spectral locus (edges) of the CIE 1931 color space as it is refracted from natural daylight. The specific spectral

location and number of stripes in rainbows appears to be potentially promising as both a guide to the general categorization of color, and as a quick functional test for dichromacy and tetrachromacy. This also enables a test of if the spectral locations of rainbow stripe boundaries are shared among dichromats, trichromats, and tetrachromats. In short, establishing color categories enables us to identify and verify which colors in a smooth gradient are most salient and prototypical based on width and center points of any perceived stripes.

Figure 1.4. Rainbows, which are comprised of a continuous change in electromagnetic frequency, can appear to have categorical stripes of color.

Past research has tested how humans categorize color according to perceptual and cognitive constraints (Boynton et al., 1989). Of particular interest in this work is parsing lexical color categories from perceptual color categories, with the emphasis being on the

latter in terms of defining color perception in a way that is free of semantic meaning, or otherwise influenced by language (Boynton et al., 1964; Boynton & Olson, 1987).

One experiment tested the difference of responses between color labels and actual color stimuli (Shepard & Cooper, 1992). When sorting colors according to labels, the outcomes were nearly identical between normal vision and colorblind participants. However, when color stimuli themselves were sorted in the same manner, colorblind individuals' rankings of colors became condensed among the red-green dimension, despite normal vision participants' responses staying the same. This work from Shepard and Cooper has been reproduced (Jameson & Komarova, 2009). These findings demonstrate the importance of using task instructions and color stimuli that are as free from semantic interpretation as possible to yield perceptually accurate data, rather than data that are more connected to language processing.

More recently, it has been discovered that color categorization begins even in infants, suggesting a biological component is responsible (Skelton et al., 2017). This is significant as previous theories attribute much of categorical color perception to communication or culture that is learned.

Dichromacy, Anomalous Trichromacy, and Tetrachromacy

Past work in color perception has demonstrated that individuals who are colorblind perceive fewer colors than those with normal vision (Gordon, 1998) and therefore upon perceiving a rainbow-like gradient are less likely to see as many stripes.

A brief overview of the major types of colorblindness is in order. Normal vision trichromats have three color cones, S, M, and L which typically correspond to 445 nm, 535 nm and 575nm for peak absorbance respectively (Fortner & Meyer, 1998). In

colorblindness, one of two deficiencies happen, the first being a reduced receptor peak absorbance curve, where any of the previously mentioned S, M, or L cones have reduced or shifted peak sensitivity (Goldstein, 2007). This is known as color anomaly. The second deficiency type is when a color cone type is completely absent, in which case only two cones function, known as dichromacy. Greek prefixes then denote precisely which of the S, M, or L cones are affected, tritos-, deuteros, and protos respectively. For instance, deuteranopia (the most common colorblind type) indicates the peak absorbance of the M color cone is skewed. Conversely, deuteranomaly denotes the complete absence of M type color cones. Using this information, informed judgements can be made as to how colorblind observers would see a typical rectilinear gradient of electromagnetic light frequencies, seeing fewer bands of color than a normal vision trichromat or having a skewed band perception depending on color vision deficiency (CVD) type.

A primary goal of this work is testing those that may have functional tetrachromacy, more often females, who are capable of discerning between more colors than trichromats (Jameson et al., 2020). In this case, it is logical to infer that tetrachromatic individuals will see more stripes of color than a normal vision participant, as has been demonstrated to a limited extent in work by Jameson et al. (2001).

Further, color researchers have posited that those with tetrachromatic capabilities may be able to process visual information more quickly due to having the extra cone receptor type (Sutherland, 2001). An potentially promising possibility with the current work is being able to use a novel technique to more efficiently identify tetrachromats than the presently time-consuming methods available such as multispectral assessment (Bochko & Jameson, 2018) or expensive, limited PCR assessment that is unable to

identify functional versus non-functional tetrachromacy and does not require expensive optical equipment (Jameson et al., 2001) and yields increased accuracy.

The current experiment tests for common categorical perception of spectral colors by having participants identify the most perceptually prominent stripes of colors they observe in a pure wavelength rainbow. Participants are shown a full spectrum gradient of color using a prism redirecting broad spectrum light directly via the sun. The goal of the experiment is to determine if there are common archetypal spectral colors expressed as dominant stripes, and if so, where these are spectrally located and how these locations correspond to models of color vision, potentially as a measure of color blindness, and tetrachromacy. Specifically, I hypothesize that those with colorblindness will perceive fewer stripes of color and those with tetrachromacy will perceive more stripes than those with typical color vision. I further hypothesize those with tetrachromatic vision will reliably identify new stripes within the green-red region. Finally, I aim to identify the prototypical colors participants see in a rainbow-like gradient, and if the boundaries of these prototypical colors are common among dichromats, trichromats and tetrachromats.

Method

Participants

Participants were 21 Arizona State University undergraduate students from the introductory Psychology course subject pool (M_{age} = 18.81, SD = 1.01; 2 African, 6 Asian, 6 Latinx, 8 Caucasian). All participants had normal or corrected to normal visual acuity and were verified to have normal color vision based on Ishihara tests (Appendix A). Special recruitment was used to include 11 self-identified color blind or color

anomalous participants (M_{age} = 19.09, SD = 1.37; 4 Asian, 3 Latinx, 4 Caucasian), which were also confirmed with the same Ishihara plates. Of these participants, 7 were deuteranopes and 4 were protanopes.

Materials

For this experiment, sunlight and a standard 6" triangular glass prism were used. The prism was affixed to an adjustable stand allowing the height to be adjusted as well as the angle of the prism itself. This allowed the light from the beam to be projected onto an 8.5x11in (21x28cm) sheet of HP Premium 32 paper with 100 brightness that was located off to the side in the shade. The rainbow from the sunlight shining through the prism was projected onto the paper and aligned with the horizontal edges of the same paper, as shown in Figure 1.5. To further standardize lighting conditions and to facilitate correct alignment, the experiment was conducted within one hour of noon on sunny well-lit days, with at least 90,000 lx verified with a REED SD-1128 photometer. To test for uniform distribution of color wavelengths in the stimulus, I used a Specim IQ hyperspectral mobile camera and took a hyperspectral high dynamic range (HDR) photograph of the refracted gradient. The resulting hyperspectral HDR photograph was then processed using Scyven software (Habili & Oorloff, 2015). This confirmed that each "band" of color in the gradient maintained the same uniform width of 2.81 in (7.14 cm), SD_{band} = 0.11in (0.28cm) over the range of wavelengths from 400nm-700nm, calibrated in steps of 2nm increments (150 steps total).

Procedure

Participants were instructed to mark on the paper where they saw boundaries in the colors of the rainbow, taking about one minute per participant, no color labels or terms were used. This allowed us to not only determine how many “stripes” they perceived but also the average perceived spectral locations and width of each stripe (Figure 1.5). In order to standardize the measurements, the prism was pointed at a tilted planar surface, aligning the edges of the gradient with the edges of the paper such that the far edges of the rainbow gradient were identically located at the same location on the sheet of paper for each participant. Finally, the lux level was measured and confirmed to be within the designated range of 97,000±1,000 Lx during each data gathering session.

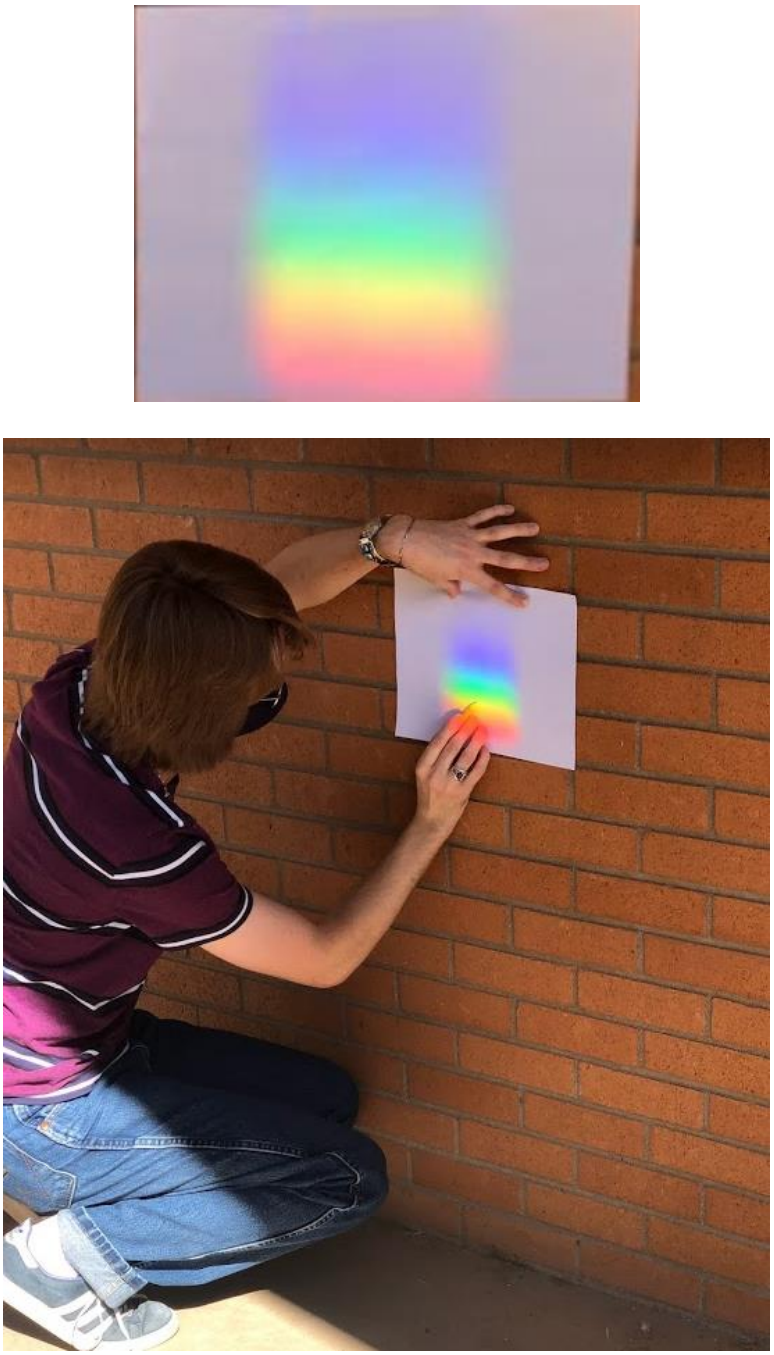

Figure 1.5. Prismatic light stimulus and participant marking boundaries. The above photos were taken using the SPECIM IQ Hyperspectral camera (top) and an iPhone X (bottom), compressed into an RGB image. This image is simulated, as participants had a clipboard behind the paper to create a smooth surface on which to mark boundaries.

Results

On average, normal vision participants were most likely to mark 6 stripes on the gradient ($\mu = 6.3$, $\sigma = 1.06$). Figures 1.6-1.7. Of the 21 normal vision participants, four saw 7 stripes, just two saw 8 stripes and only one participant perceived 9 stripes. In contrast, colorblind participants saw significantly less stripes ($\mu = 4.45$, $\sigma = 0.52$) than normal vision participants, t (30) = 5.49, p = 001, d = 1.00, a large effect size.

In order to test if the locations of the chosen demarcation line placements between color band stripes occurred at common spectral frequencies, a k-means cluster analysis was conducted. Here each demarcation point was coded by computing the physical distance between that point and the group center of apparent clusters, and then was classified to be in the group closest to that point (Wagstaff et al., 2001). The group centers were then recomputed by taking the mean total of the vectors in the group, in essence clustering points to the nearest centroid, and transforming physical distances into the corresponding electromagnetic spectral frequency. This cluster analysis revealed that average demarcation line placement among the normal vision participants began at 682nm then lines were placed at 639nm (1st stripe), 600nm (2nd stripe), 562nm (3rd stripe), 478nm (4th stripe), and 427nm (5th stripe), Figure 1.8. Participants who perceived more than 5 color demarcation lines saw additional lines placed at 522nm (6th stripe), 615nm (7th stripe), 655nm (8th stripe) and 448nm (9th stripe).

Colorblind individuals placed their beginning stripe at 672nm, then placed lines at 616nm (1st stripe), 557nm (2nd stripe), 513nm (3rd stripe), 451nm (4th stripe), Figure 1.9. A bit less than half of the colorblind participants saw an additional 5th stripe placed at 595nm. Further, colorblind participants were significantly more likely to place their

beginning stripe much lower in the spectrum than normal vision participants, t (30) = 3.14, p = 0.004, d = 0.57, a medium effect size.

Pairwise comparison tests were assessed between the normal vision and colorblind individuals, finding all but 11 of the possible 50 pairings significantly different ($p < .001$), Figure 1.10. This pointed to the possibility that normal vision participants and colorblind participants categorically perceived similar category boundaries in these nonsignificant pairings. To test if this was the case, equivalency testing was performed (Harris et al., 2012; Lakens et al., 2018). Equivalency testing calculates the confidence interval between means of a given pair of samples and then checks to see if the confidence interval completely falls within the equivalence threshold, based on effect size, $-d$ to d where $d = (M_2 - M_1) / SD_{pooled}$. If the confidence interval falls within the equivalence threshold it is assumed that the pairing is equivalent. Following the advice of Harris and colleagues (2012), a confidence interval level of $(1- 2\alpha)$ was used, resulting in 5 of the pairings being equivalent, Figure 1.10, Table 1.1.

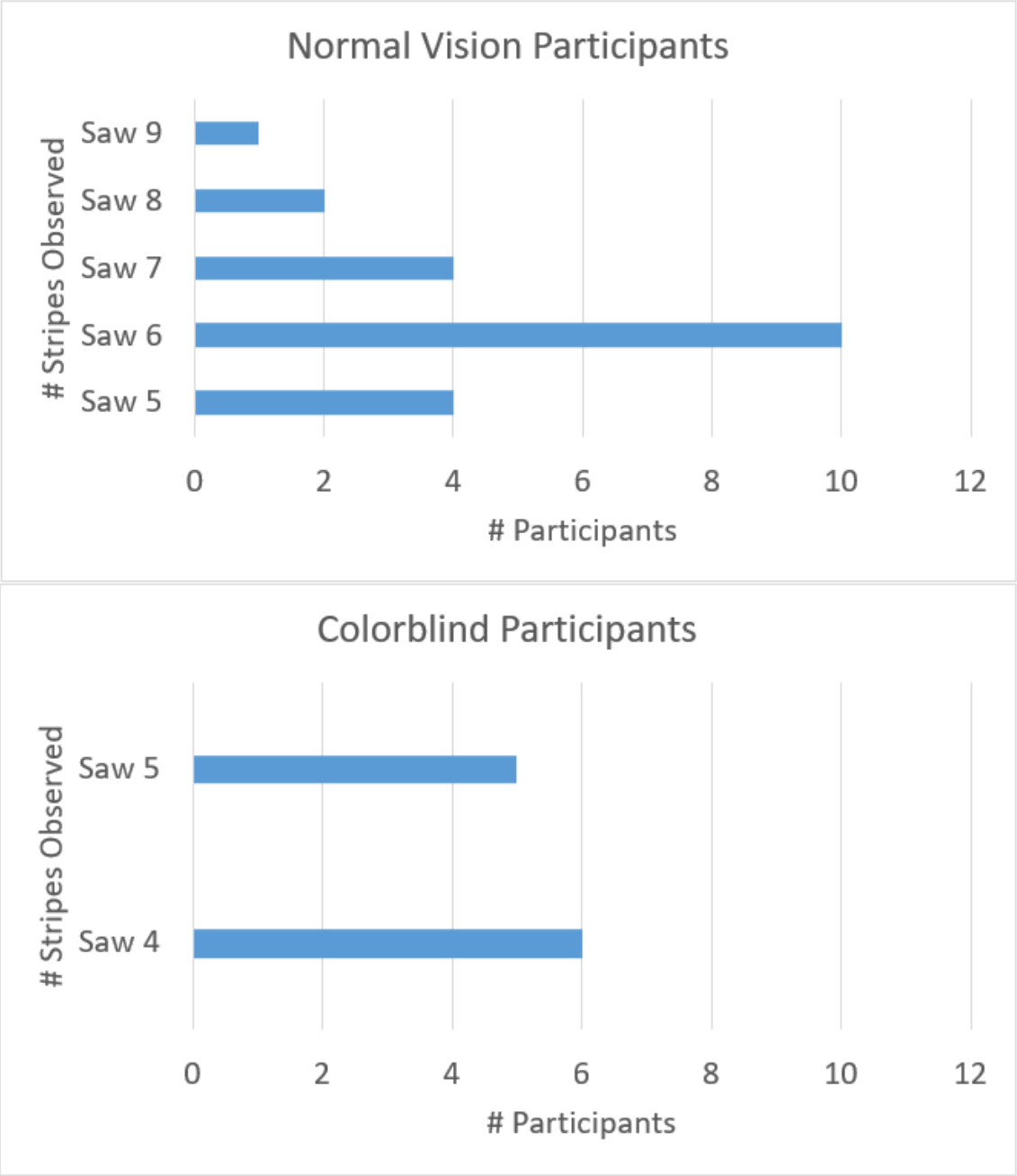

Figure 1.6. Graphs depicting how many stripes each participant saw in the normal vision group (top) compared to colorblind participants (bottom). Most normal vision participants perceived 6 stripes, and color blind participants identified significantly fewer stripes than the normal vision participants.

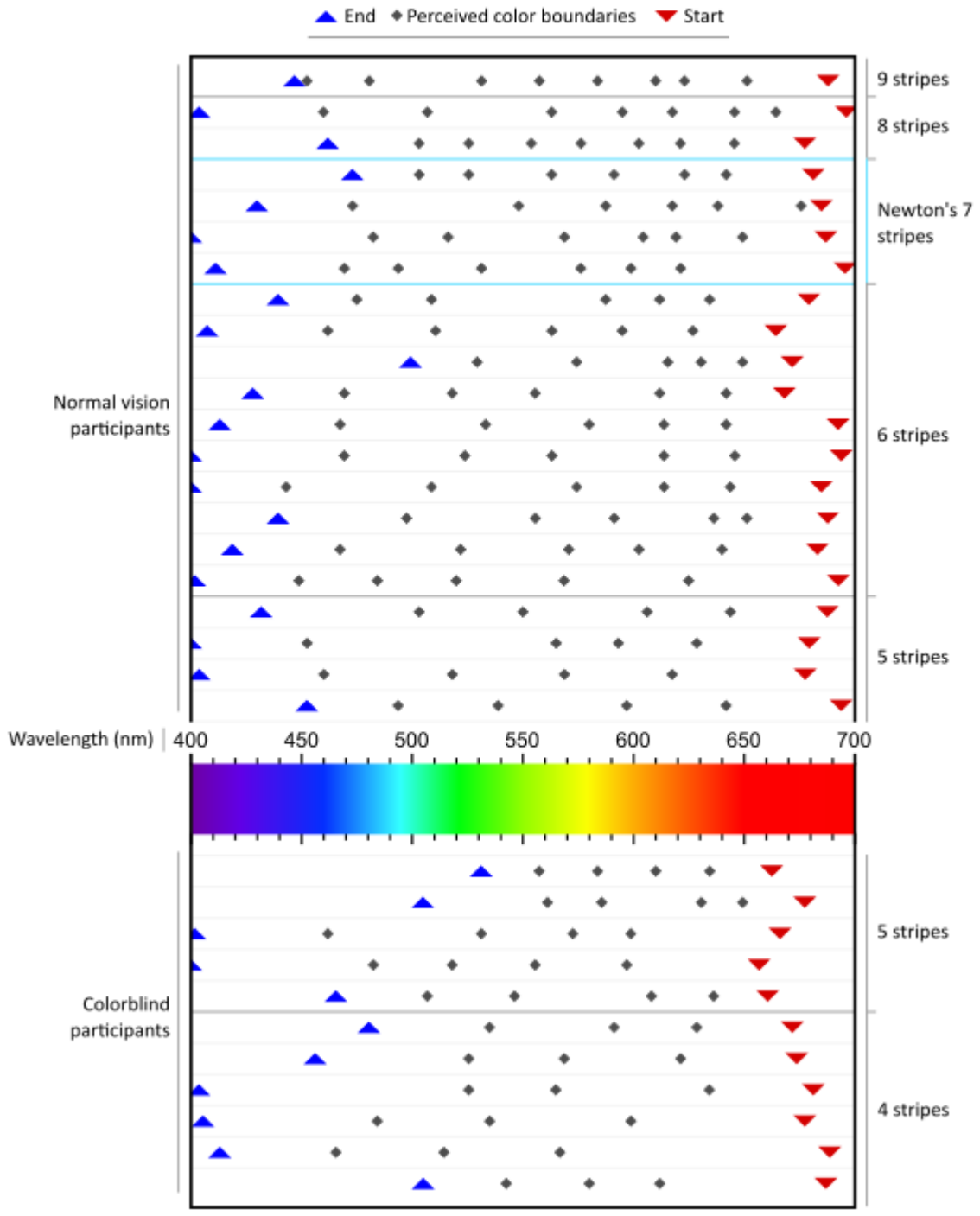

Figure 1.7. Graph indicating where participants placed markers between color boundaries. The top indicates participants with normal trichromatic vision whereas the bottom portion shows data from colorblind participants.

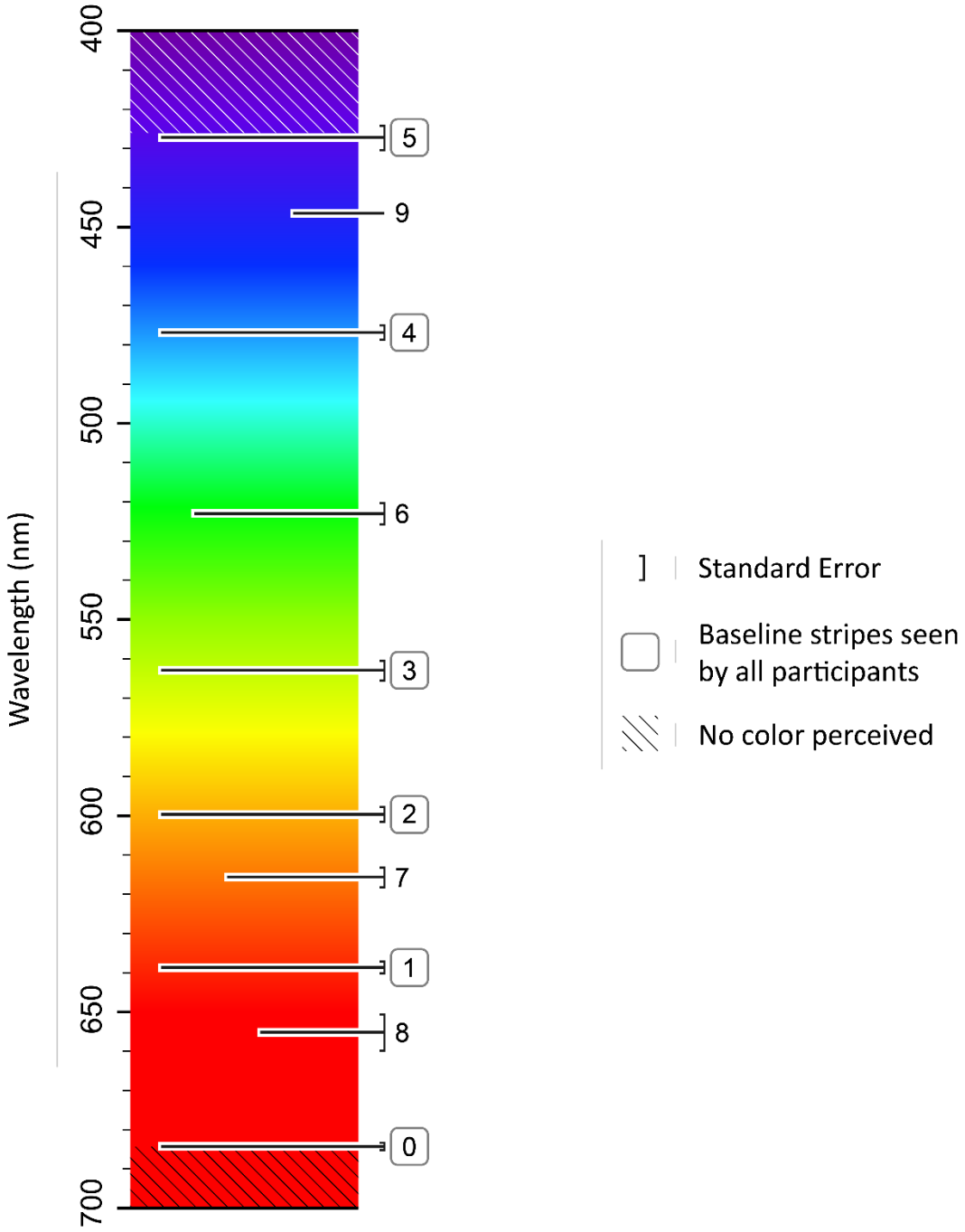

Figure 1.8. Cluster analysis reveals where normal vision participants are most likely to place line demarcations. 0 represents where participants indicated they first started to perceive color. Numbers 1-5 correspond to line placement for all participants. Participants that perceived more than 5 stripes have placed new stripes at locations 6, 7, 8, and 9 respectively. Note that those who perceived 7 and 8 stripes had increased differentiation in the orange-red range with new stripes located at 615nm and 655nm. No standard error is available for the 9th stripe as only one participant marked 9 stripes.

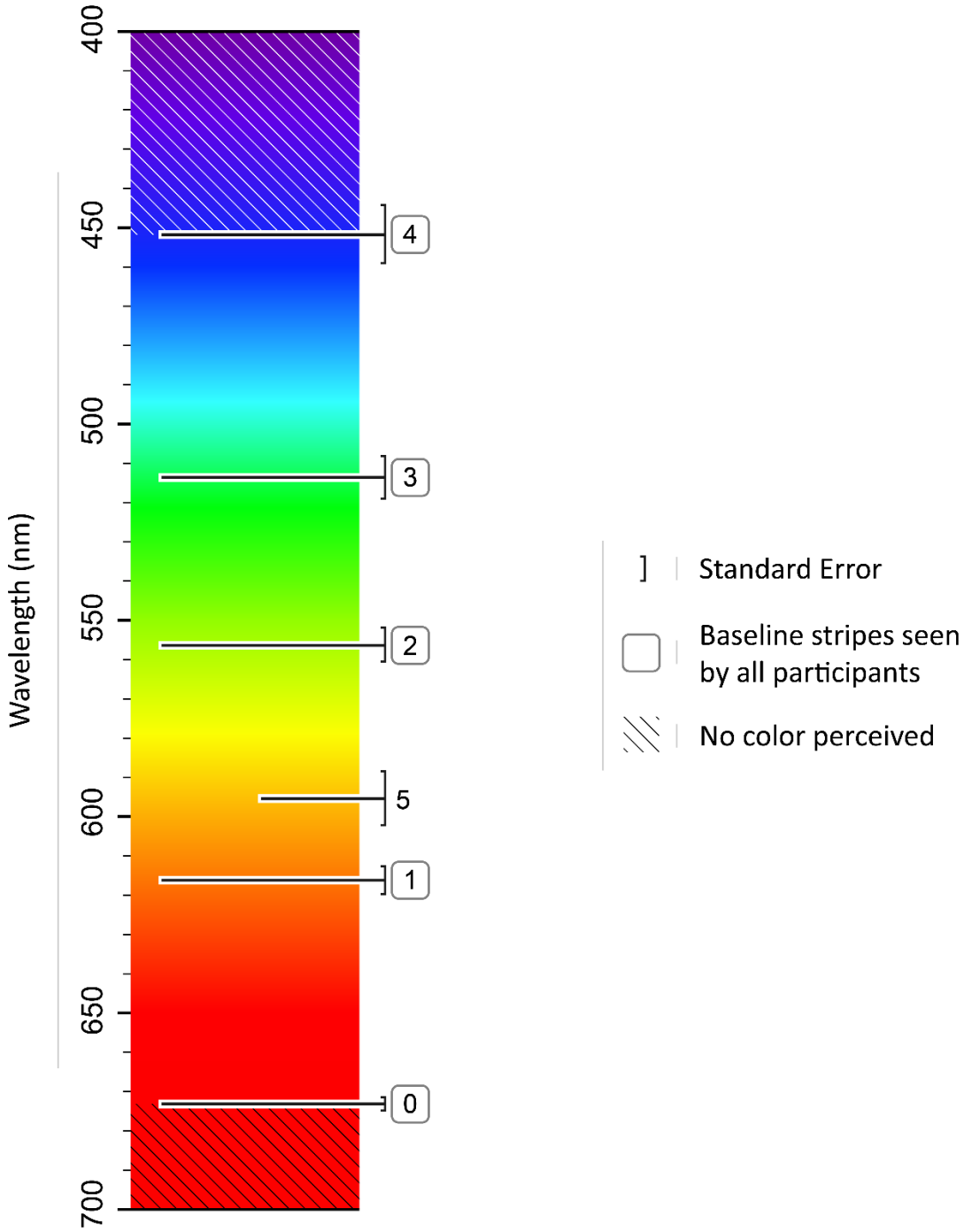

Figure 1.9. Cluster analysis shows that colorblind participants place stripes demonstrating a collapse amongst the red-orange dimension when compared to normal vision participants.

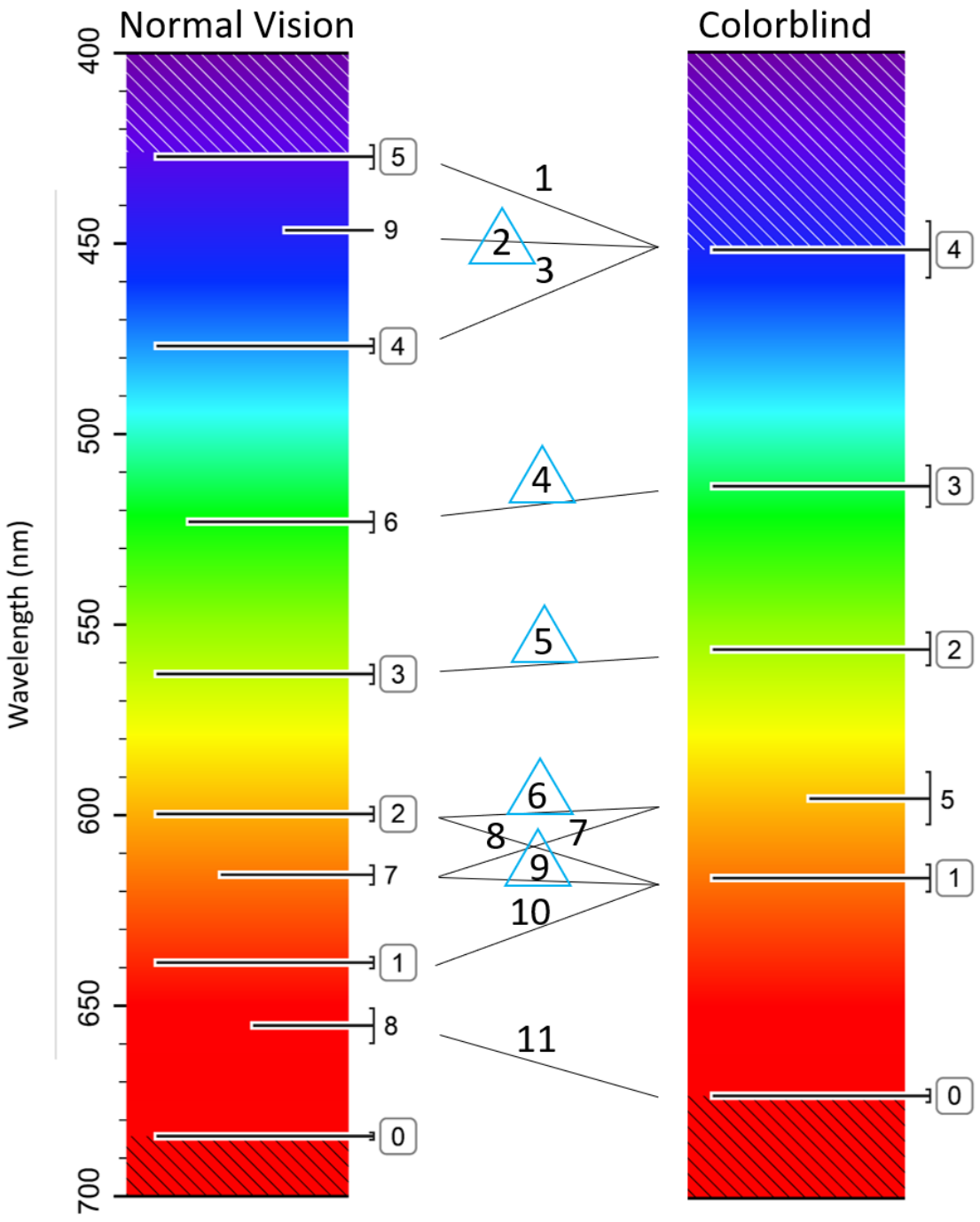

Figure 1.10. Comparison of non-significant pairwise tests between line demarcations (numbered lines between spectra). All other pairings were significantly different ($p < .001$). Blue triangles represent equivalent pairings. These pairings are further analyzed in Table 1.

Table 1.1.
Equivalence Tests Between Color Demarcation Pairings

Pairing	Confidence Interval	*-d to d*	*p*	Equivalent?
1	-0.07 - 1.42	-0.68 to 0.68	0.14	N
2	-0.34 - 0.15	-0.40 to 0.40	0.92	Y
3	-0.14 - 1.35	-0.76 to 0.76	0.13	N
4	-0.23 - 0.29	-0.33 to 0.33	0.45	Y
5	-0.19 - 0.17	-0.24 to 0.24	0.49	Y
6	-0.17 - 0.18	-0.21 to 0.21	0.78	Y
7	-0.28 - 2.36	-0.32 to 0.32	0.22	N
8	-0.09 - 1.60	-0.84 to 0.84	0.09	N
9	-0.32 - 0.37	-0.42 to 0.42	0.84	Y
10	-0.05 - 2.06	-0.71 to 0.71	0.13	N
11	-0.05 - 2.32	-0.86 to 0.86	0.09	N

Discussion

This experiment looked at how observers categorized number and width of stripes after observing a naturally-produced gradient of light wavelengths. Unlike the prototypical ROYGBIV rainbow that contains seven stripes of color, on average participants perceived just over six stripes, roughly corresponding to red, orange, yellow, green, blue and violet. Moreover, viewers with normal trichromatic vision saw significantly more stripes than dichromatic color blind individuals. While at this juncture in time I am unable to test individuals who are believed to be tetrachromatic, these results point to some of the participants, those that perceived more than 7 stripes, to potentially be tetrachromatic, especially the one that saw 9 stripes. The results of the cluster analysis support the notion that tetrachromatic individuals likely possess increased sensitivity and color discriminability around the spectral frequency extremes rather than the central frequencies around 480-660nm. When compared to normal vision participants who

perceived 6 stripes, colorblind participants' locations of stripes were similar for placement of their middle point markers in the lime-green and blue-green regions of the stimulus but toward the ends of the distribution, colorblind observers showed a collapsing of their red-orange and violet-blue markers. Consistent with past work on colorblind sensitivities, colorblind individuals also placed their beginning marks at shorter wavelengths than color vision normal participants, indicating decreased sensitivity in the red region of the spectrum.

Categorization of color between the normal vision trichromats and colorblind participants showed common category boundaries around blue, green, yellow, and orange stripes. Because of colorblind individuals having decreased sensitivity to red, especially protanopes (Fortner & Meyer, 1998), there was not an equivalent pairing around the red stripe, with normal vision participants placing their initial red line much farther toward the longer wavelengths as expected. Additionally, whereas normal vision participants differentiated between blue and violet, colorblind individuals did not. This consistency is aligned with previous work showing that within perceptual spaces, observers tend to respond similarly (Ashby & Gott, 1988).

From the standpoint of the smooth continuum of electromagnetic spectral frequencies, color categories could be thought of as a type of illusion, and while the color categories might simply be explainable due to the limited number of cone types, they also may reflect some type of evolutionarily selected functional differentiation such as generically discriminating bluish liquids, greenish plants, and reddish earth tones, Figure 1.11. It would appear the human visual system is optimized for leaf and fruit consumption, where discriminating between red and green is critical (Gegenfurtner &

Rieger, 2000; Lucas et al., 2003). More precisely, the color of leaves serves as an indicator of maximum nutritional content with leaves changing between green and red. Further, most fruits generally begin as green but ripen to become yellow, orange, brown or red, hence being "red-shifted" (Willson & Whelan, 1990) other research points to an additional benefit as far-away fruits in this color region are easier to identify against green leaves (Bompas et al., 2013). This helps explain why the human visual color system has increased discrimination in the red-green portion of the spectrum, and also bolsters the results that increased discernment between spectra "stripes" occurs in the red-green portion of the color spectrum.

Figure 1.11. Examples of categorical colors for which the human visual system may be optimized. These categories map onto the perceptual categories chosen by participants in this experiment.

Other evidence exists to support M and L cones being most optimal for human vision and hence their increased range of sensitivity in the red-green region (Changzi et al., 2006). Blood oxygen saturation is readily visible in skin colors of all humans and humans have an incredible sensitivity to skin tone changes as a function of blood oxygenation. This is important for mate selection as many underlying health issues will impact blood oxygenation levels. In addition, being discerning of healthy skin tone aids in discernment of emotional states in others. It is theorized that this is why humans and other primates are bare faced. Further, particular colors of commonly encountered objects seem to be grounded in interesting phenomena related to why we see color. For instance, dirt comes in four main colors: white, which is influenced by salt deposits; brown, influenced by decomposing plant material; yellow or red, revealing presence of iron; and gray, which signals water has been in the soil for a long period of time (NRCS, 2021). Green in plant leaves is attributed to the two main types of chlorophyll which both absorb violet and orange and blue and yellow light respectively, while absorbing no green, but subsequently reflecting green (Milne et al., 2014). Another fundamental color seen in nature is blue in water which is attributed to a combination of the reflection of the blue sky and the absorption spectra of water, being highly contingent on observation angle (Braun & Smirnov, 1993). Encounters with these and other common colors might influence the way that perceptual color categories are formed as increased experience with color stimuli helps to develop classification of the same categories (Homa et al., 1981).

The real-world continuous color stimulus gradient and line demarcation task can also provide fruitful results and methodology for those interested in categorical color

perception. Previous research into color categorization often entails participants matching or locating semantic color labels (Uchikawa & Shinoda, 1996) whereas the current research model presents a novel approach to locate color boundaries without need of explicit color naming. Further, as pointed out by Bae et al. (2015) past work in categorical color perception has been limited by two methodological obstacles. The first is that previous research has used stimuli encompassing a restricted area of color space and the second issue regards the sampling resolution: if color stimuli are quantized too coarsely, then it could lead to the incorrect conclusion of exact discrete boundaries. The current stimulus is a full set of pure color wavelengths that ranges from 400-700nm with a high amount of resolution and the task itself allows participants to place their responses at any location and make as few or many marks as desired, thus mitigating any confounds related to stimulus resolution or quantization of participant responses.

Another level of explanation that would provide mechanisms of these results may be that categorical boundaries were formed from the overlapping sensitivities in the color cones of the human retina, Figure 1.12 (Fortner & Meyer, 1998). Given how each cone can be stimulated in unique ways from pure wavelengths (Figure 1.13), it is possible unique pairings may result in categories based on single color cones being excited or unique combinations of color cones to create a categorical perception of color when viewing a continuous rainbow-gradient stimulus. Alternatively, the stripes could be a function of peak sensitivities in combination with points of equal overlap between adjacent cones (Figure 1.14).

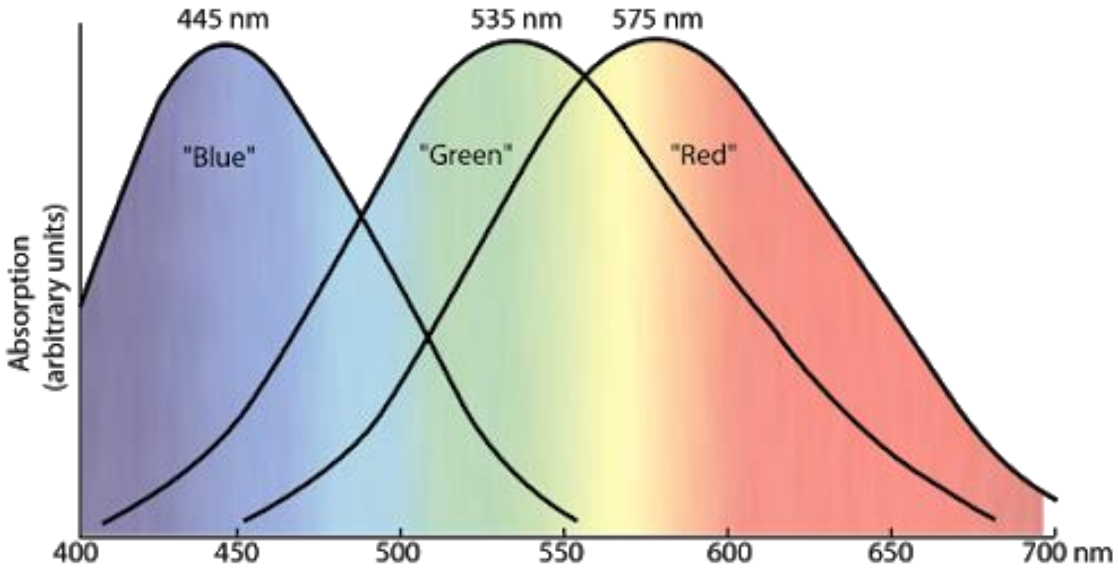

Figure 1.12. Absorption rates of color cones of human eye.

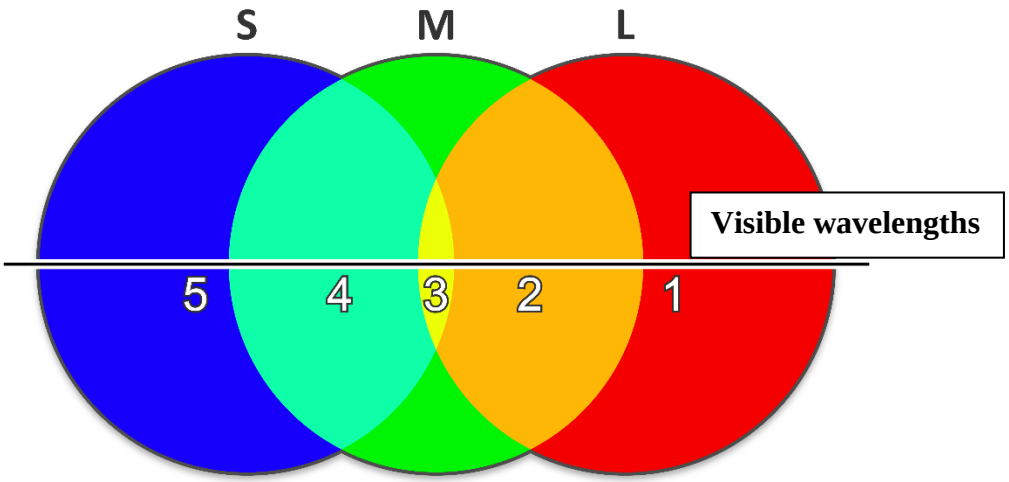

Figure 1.13. Conceptual diagram of how participants may mark boundaries. Each of the color cones and their interactions with the other cones, S, M, and L may be responsible for perceived stripes seen in the smooth rainbow-like gradient. The line in the middle of the diagram represents wavelengths visible to the human eye.

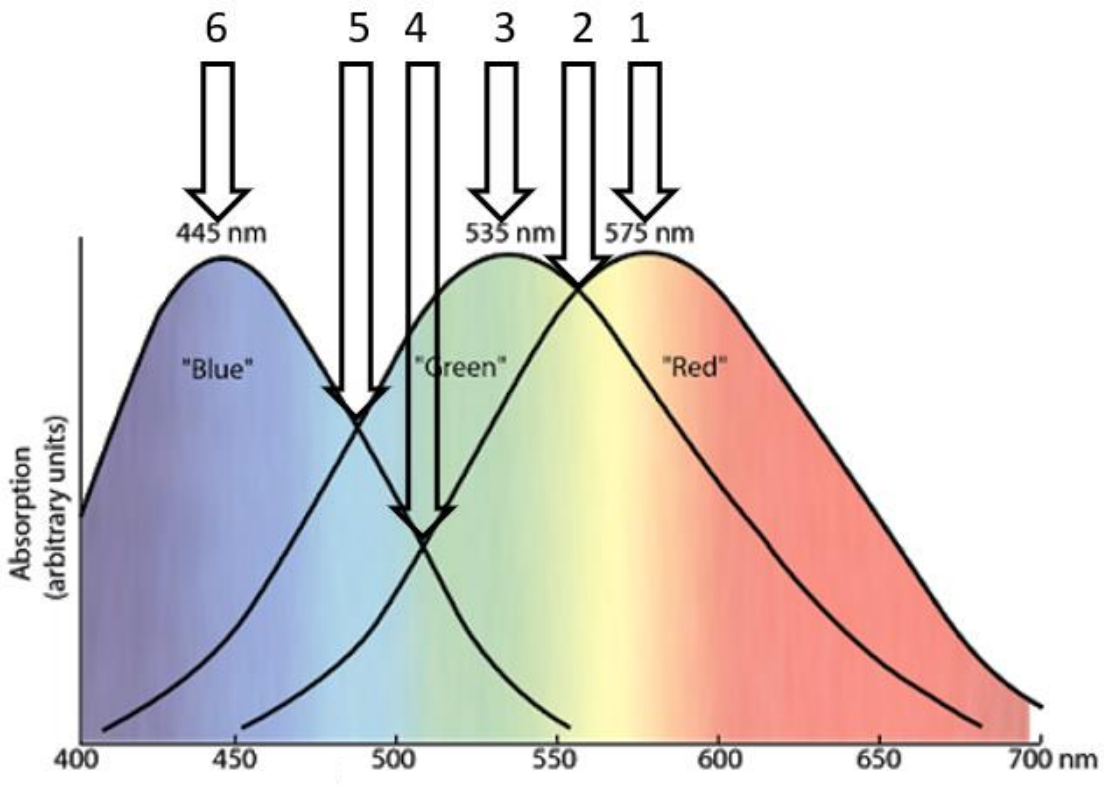

Figure 1.14. An alternative conceptual model for seeing stripes in smooth rainbow-like gradients. Stripes in this case are seen at peaks of S, M and L cone absorption curves as well as where the curves equally overlap.

The current study supports that tetrachromatic individuals see more “stripes” than trichromatic individuals, and importantly that some of the increased discernment may happen at the 615nm and 655nm ranges where new stripes emerge. Using the current rainbow stripe counting method with an inexpensive prism, allows functional tetrachromacy to be tested more widely to obtain precise prevalence of this unusual exceptional trait as well as expand on current methodologies for assessing categorical color perception.

CHAPTER 2

THE TEMPORAL CLOCKING RATES OF WHERE (DORSAL) AND WHAT (VENTRAL) VISUAL SYSTEMS: MEASUREMENT OF MOTION AND COLOR-HUE DIRECTION

In everyday experience, color is not only perceived in a static manner, but commonly there are temporal variations from motion and changes in lighting conditions as observers and objects move about in the world. Much of the research done on dynamic color perception in the past decades has focused on color constancy under a variety of largely static settings (Maloney & Wandell, 1986; Witzel & Gegenfurtner, 2018), while ignoring temporal properties of color perception that may now be fruitfully studied with new fast-display technology that has become available. The current research explores temporal properties of color vision that may be functional in ways that can enhance perception in new display designs, and further quantify processing limits of the visual perception system.

Goodale and Milner (1992) describe a dual stream brain processing model, characterized as the dorsal and ventral streams, which are respectively responsible for "how/where" versus "what" types of decisions based on original work from Mishkin and Ungerleider (1982). This model has been found to apply to the temporal perception of color in both the speed that a colored object is moving (where) as well as the speed of discerning hue changes (what), (Norman, 2002). Specifically, visual *where* decisions appear to require less cognitive effort, are focused on motion and largely include inputs that vary across large portions of the retina. Further, such decisions are sensitive to low spatial frequency and high temporal frequencies (Merigan, et al., 1991). In contrast,

visual *what* decisions appear to utilize higher cognitive effort, are focused on details and are primarily foveal while being sensitive to high spatial frequency and lower temporal frequencies. In both cases objective experiments can be carried out to measure the processing rate of the how/where or what systems by determining the point at which an observer can no longer accurately determine specific aspects of color stimuli. Currently, there is some research linking the two-streams paradigm to critical flicker fusion rate (Holloway et al., 2012; 2013), but little specifically testing aspects of color, and hence this is a potentially valuable avenue to explore. Determining processing rates can further elucidate the temporal information processing limits of the brain using color stimuli that is largely free from semantic interpretation, in contrast to previous research using representational stimuli like words, scenes and shapes (Marois & Ivanoff, 2005; Shapiro, 2001).

A useful and traditional method for testing visual temporal perception is the critical flicker fusion (CFF) task which yields precise thresholds for temporal limits (Eisen-Enosh et al., 2017). An extension of this task has appeared in work by Holloway and McBeath (2013) in which two tasks were devised to separate visual processing by the *what* and *where* visual processing streams and then compared to CFF. The first task was one of apparent motion where a set of green LEDs moved in clockwise or counterclockwise motion which allowed for an objective measure of the threshold for a *where* processing system. A second task comprised of a set of shapes either increasing or decreasing in number of sides to objectively measure the *what* system threshold. Performance on these two tasks was compared to CFF rates in the same participants, demonstrating that the *where* task correlated well with CFF and produced similar

frequency thresholds whereas the temporal threshold rates of the *what* task were slower as is expected of a cortical system that requires more conscious cognitive processing. The methodology of this task can be extended into the realm of color perception by using several colored monochromatic LEDs to test color perceptual thresholds in a *where* system while identification of order of serially flashed monochromatic LEDs in the same location could be used to test visual thresholds of a *what* processing system.

Past work suggests differentiation of CFF temporal thresholds by wavelength (Green, 1969). Long (L) and medium (M) wavelength color cones tend to have similar CFF thresholds to each other around 20Hz whereas small (S) color cones have lower CFF thresholds around 15Hz. However, it remains unknown if apparent motion is impacted by individual wavelengths and by extension the limits of the *where* processing system.

Some past research has attempted to define temporal color processing limits using a paradigm called critical color flicker fusion (Nakano & Kaiser, 1992; Truss, 1955). Critical color flicker fusion entails flashing two lights alternately and increasing the frequency until it appears the two colors have combined to form a new color, for example flashing a red light and a green light alternately fuse to produce a yellow color. Critical color flicker fusion differs from traditional CFF in that the former happens at a slower rate than the latter while involving an intentional chromatic component; once the two colors have fused, flickering still appears until the frequency rate is turned up high enough that they fuse to create a steady light as seen in CFF. Findings within this realm have been mixed, however, likely due to small sample sizes. For instance, Truss (1955) finds that various colored pairs produce differing thresholds between two participants (Figure 2.1) whereas Nakano and Kaiser (1992) find no difference between pairs of

colors between four participants (Figure 2.2). From the figures, one can also see the limited pairings of colors being displayed to participants, indeed these colors were created using glass color filters that are known to have a wider band spectrum per color than modern LED equivalents with narrower, precise spectra.

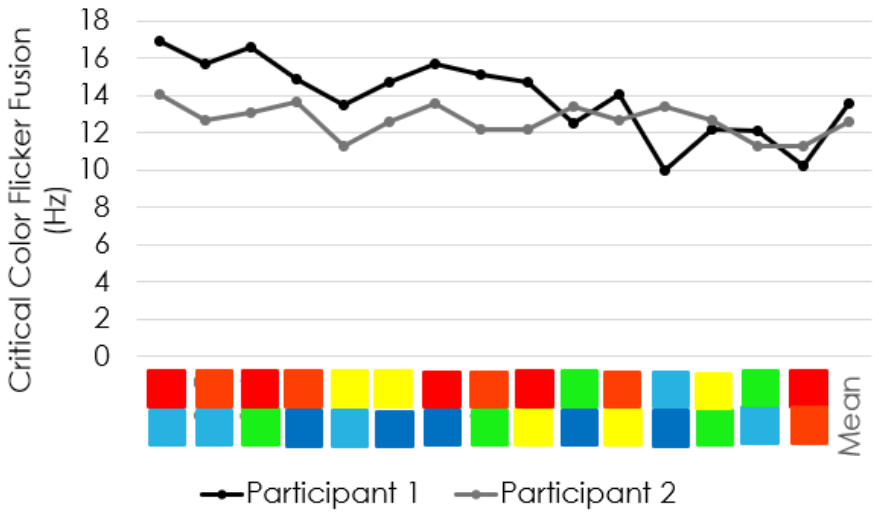

Figure 2.1. Critical color flicker fusion rates reproduced from Truss (1955). Significant variability exists between each participant based on which color pair is presented.

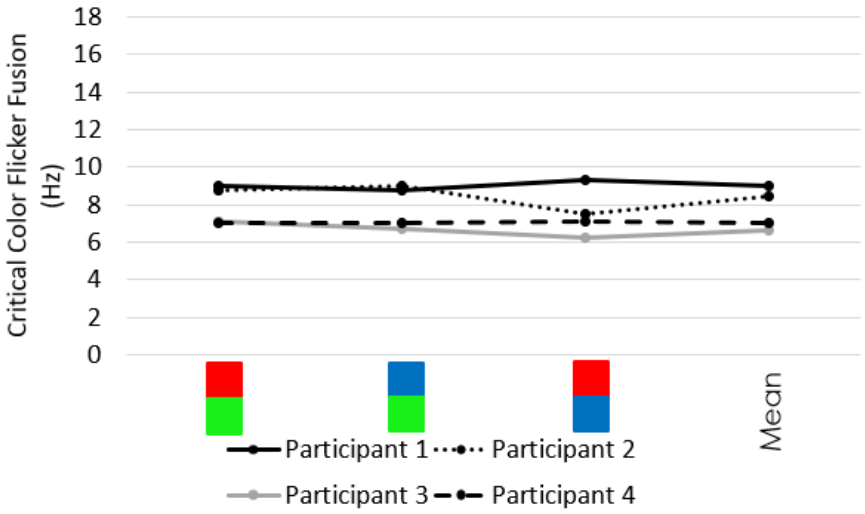

Figure 2.2. Critical color flicker fusion rates reproduced from Nakano and Kaiser (1992). Participants show little variability in their fusion rates between pairs of colors as well as an overall lower threshold (Hz) in where colors fuse.

Color research in the past century has improved a great deal with more precise measurement methods and better equipment. However, overall, much of the equipment previously used has not allowed as precise temporal and spectral control over various wavelengths as is now available, with past research typically relying on lighting sources based on the fluorescence of materials such as neon, argon, or tungsten (Campenhousen et al., 1992) or color filters with relatively wider band spectral characteristics (Sill, 1961). More recent work has begun to utilize newly developed narrow spectra LEDs for color vision experiments, which enable a higher degree of control over intensity, wavelength, and temporal frequency (Finlayson et al., 2014). For example, a recent study showed that the duration of blue is overestimated when compared to red, which warrants further investigation into how colors are temporally perceived (Thones et al., 2018) and others have demonstrated how perceived color transitions between primary colors are nonlinear despite being modeled as perceptually linear in CIELAB space (Kong et al., 2019).

The current experiment uses precise frequency and timing LED technology to test the perceptual processing thresholds of the perception-action based *where* system and the perceptual processing limits of the conscious-cognitive based *what* system.

Ventral tasks are usually thought to be conscious decisions that require a higher level of processing, which typically takes more time (McBeath, Tang, Shaffer, 2018; Norman, 2002), and therefore it is hypothesized that participants will perceive order of color in the *what* task more slowly (a lower CFF threshold) than the motion of color in the *where* task due to the former process requiring conscious cognitive processes versus a more automatic perception-action based process in the latter process. Further, I expect a

high level of precision of these two systems in determining exact thresholds based on objective responses rather than subjective responses of past work.

Method

Participants

Participants were 10 undergraduate students at Arizona State University (M_{age} = 21.22, *SD* = 2.13; 6 females, 4 males). Participants were screened for normal color vision using Ishihara test plates.

Materials

To produce the colored stimuli for each experiment, I used a custom "HyLighter" device comprising of a cluster of 10 LEDs (Figures 2.3-2.4) connected to an Arduino Uno Rev3 board. The Arduino board was in turn connected to an Adafruit TLC59711 driver board for precise control over individual diodes. Each of the 10 narrow-band LEDs covers a different portion of the visible color spectrum ranging from 440nm to 740nm (Figures 2.5-2.6). This device can display a broad spectrum of color and offers a high degree of control over traditional trichromatic color displays. Participants were placed in a dark room and with the aid of a chin rest, sat two feet (60cm) back from the array of LEDs in each experiment. This enabled us to present stimuli that were at an 8° paracentral visual field view for maximal stimulation of color cones. All LEDs were tuned to emit 150 lux such that participants could not use brightness as a cue for color direction; a REED SD-1128 photometer was used to verify uniform lux output between LEDs.

In the *where processing* experiment, I placed four HyLighter devices in a diamond array, and used six prototypical colors as defined by previous work on color categorization (Krynen & McBeath, 2021). Specifically, the six LED colors used were: Violet (450nm), Blue (473nm), Green (530nm), Yellow (587nm), Orange (613nm), and Red (660nm). The individual LEDs were switched on sequentially to give the viewer the perception of motion moving either in the clockwise or counter-clockwise direction continuously until the participant responded (Figure 2.7).

Figure 2.3. Layout of HyLighter board. Individual clusters of 10 LEDs are shown switched off.

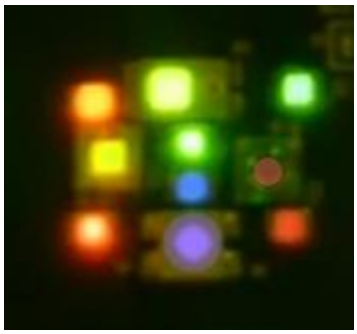

Figure 2.4. Photo of cluster of LEDs with all diodes switched on to give a sense of the actual color output.

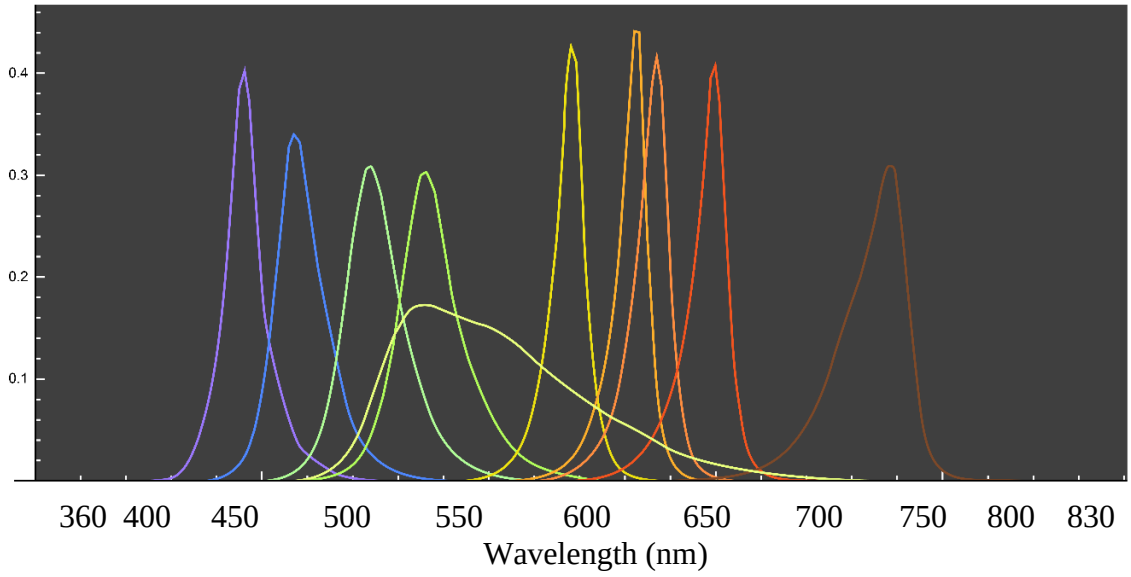

Figure 2.5. Spectral output by individual LED. Note that the majority of LEDs emit narrow spectra for more precise color control.

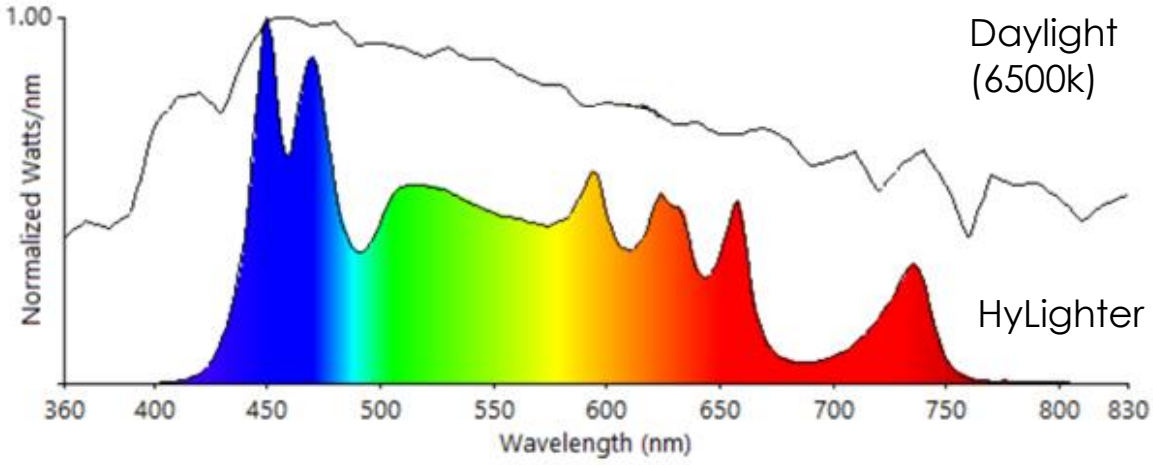

Figure 2.6. Combined spectral output of all LEDs compared to daylight (top curve in figure). When tuned, spectral output is similar to daylight peaks.

For the *what processing* experiment, I used a single "HyLighter" device which had a 10in (25.4cm) diameter optical tube installed on the top of the LED board to focus the light onto a single point. The reason for using this optical tube was to remove any

possible spatial information a participant could associate with each LED as it illuminates. In this experiment, order of color was assessed between blue (473nm), green (530nm), and red (660nm) LEDs. The LEDs were programmed to illuminate continuously in a sequential Blue-Green-Red or Red-Green-Blue order until the participant responded (Figure 2.8).

Procedure

The experimental task in both experiments utilized the method of motion handedness (McBeath, 1990) in which participants do not simply make a subjective judgment of when visual items reach a set threshold, but rather make an objective judgment of which *direction* the color appears to travel, either clockwise or counter-clockwise, or in the second experiment between order of Blue-Green-Red or Red-Green-Blue. Using this method is superior to previous methods of testing flicker rates because each trial has an outcome that can be scored as correct or incorrect. In other words, the method of motion-handedness is the illusion of motion with an objective direction, compared to other test methods that rely on the subjective claims of fusion by the participant.

A staircase procedure was used to determine the threshold, measured in Hz, at which participants can no longer distinguish the physical direction of color movement (Levitt, 1971). The adaptive staircase method remains a superior test of CFF due to its reliability and minimization of time compared to other popular psychophysical procedures like the method of constant stimuli and the method of limits (Eisen-Enosh et al., 2017).

Following McBeath (1990), if a viewer answered incorrectly for motion direction, speed of presentation was reduced in the next trial. Speed remained the same on the next trial if the participant answered correctly on the last trial or last two trials previous to a correct trial and finally, speed increased if three trials in a row were identified correctly. The experiment terminated when the participant experienced nine reversals (of speeding up and slowing down). By using these thresholds, I was also able to control for guessing. Each experiment began at 1Hz and was changed by incremental intervals of 20ms in the first block, then 10ms in the second block and finally by 5ms in the final block. Each block began at the speed where the previous block concluded, allowing discovery of precise thresholds at which participants were able to successfully discriminate direction of color. Overall, participants completed each experiment in about 20 minutes for a total average of about 40 minutes to complete both the *where* and *what* tasks. An example trial would consist of the stimulus moving at a speed of 1Hz in both tasks, with the participant being instructed to respond direction (clockwise/counterclockwise) or order (red-green-blue/green-blue-red) depending on the experiment, and then the speed of the subsequent trial would change based on the aforementioned staircase procedure until nine reversals were observed. Once a trial block concluded, a new trial block with finer thresholds was used, until all three blocks (20ms, 10ms, then 5ms) were presented and participant responses were recorded, and then the next experiment would begin. The *where* and *what* experiments were counterbalanced such that half of participants started with the *where* experiment and the other half with the *what* experiment, then completed the other paired experiment.

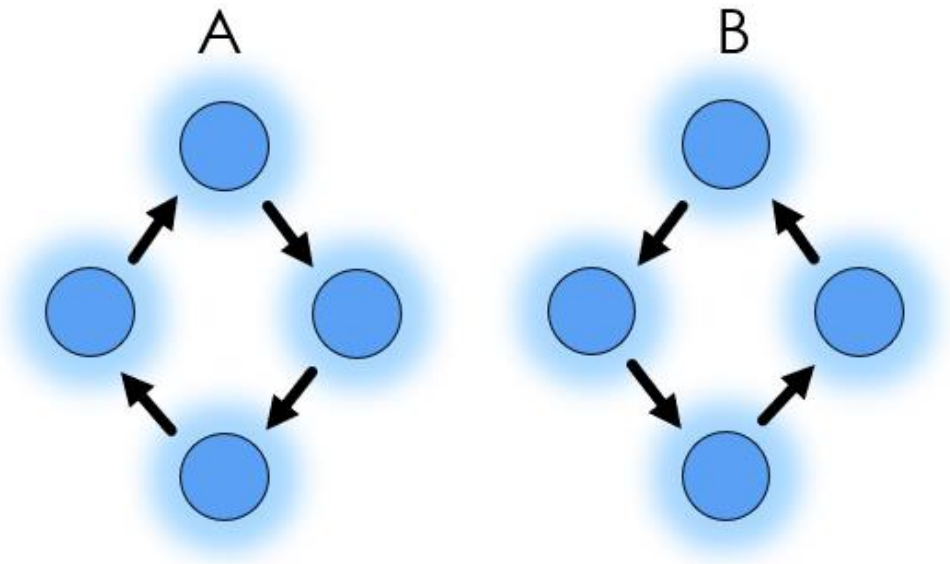

Figure 2.7. Orientation of blue wavelength LEDs using method of motion handedness in Experiment 1. By sequentially illuminating LEDs in a diamond-like pattern, an observer can evaluate perceived direction and hence assess a correct/incorrect response. When the LED's flicker too fast, participant can no longer ascertain the direction of travel. The threshold of when participants can just barely discern direction of travel indicates the measure of maximum perceived flicker rate.

The thresholds found are not expected to precisely model actual color cone receptor flicker fusion rates, but rather provide a behavioral measure of overall flicker fusion rates at different spectral frequencies, centered on the prototypical values determined in Krynen & McBeath (in progress).

In experiment 1, the LED pattern continued to flash in clockwise or counter-clockwise fashion until the participant pressed the arrow keys "→" or "←" to indicate clockwise or counter-clockwise motion respectively. Once the participant responded, the

trial terminated and the next trial began, with the actual direction of motion being chosen at random and according to the staircase procedure mentioned previously.

In the second experiment, participants specified if the colors were presented in Blue-Green- Red or Red-Green-Blue order continuously. While the pattern was repeating, participants had two clickable graphic options presented to them on a small computer display showing two bands of colors, a band in Blue-Green-Red order, and one that showed the colors displayed in the opposite direction, Red-Green-Blue to simulate the two possible color directions a participant might perceive in each trial (Figure 2.9). Once the participants clicked on a band, the next trial began, with a randomly selected sequence of Blue-Green-Red or Red-Green-Blue.

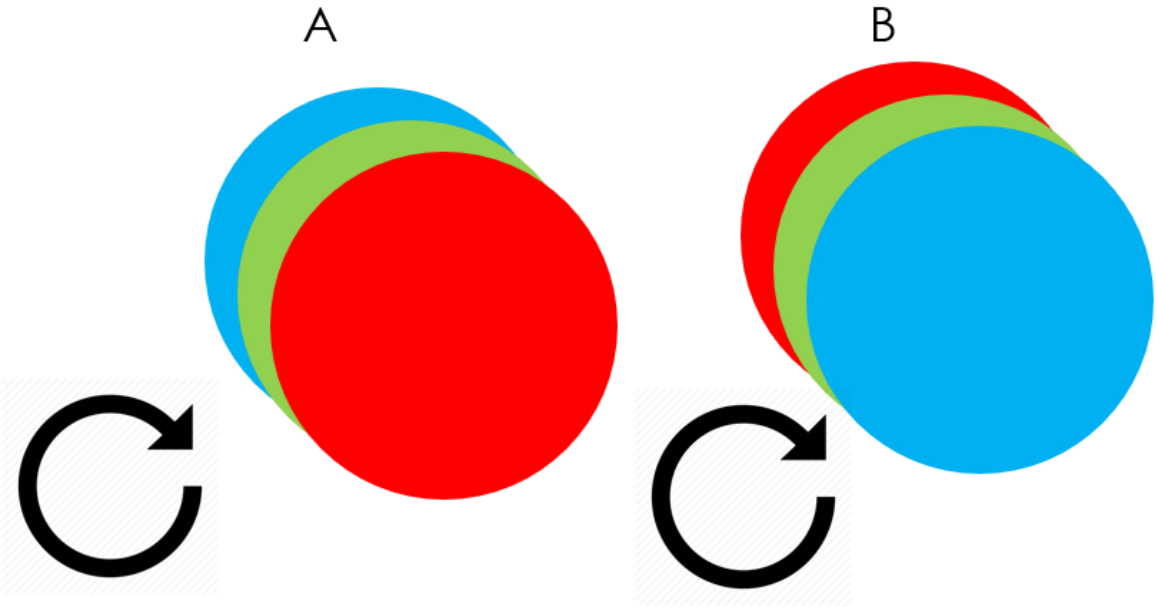

Figure 2.8. Sequence possibilities of colors for Experiment 2. Panel A shows colors moving in a "clockwise" chroma direction (Blue-Green-Red) and panel B the opposite (Red-Green-Blue).

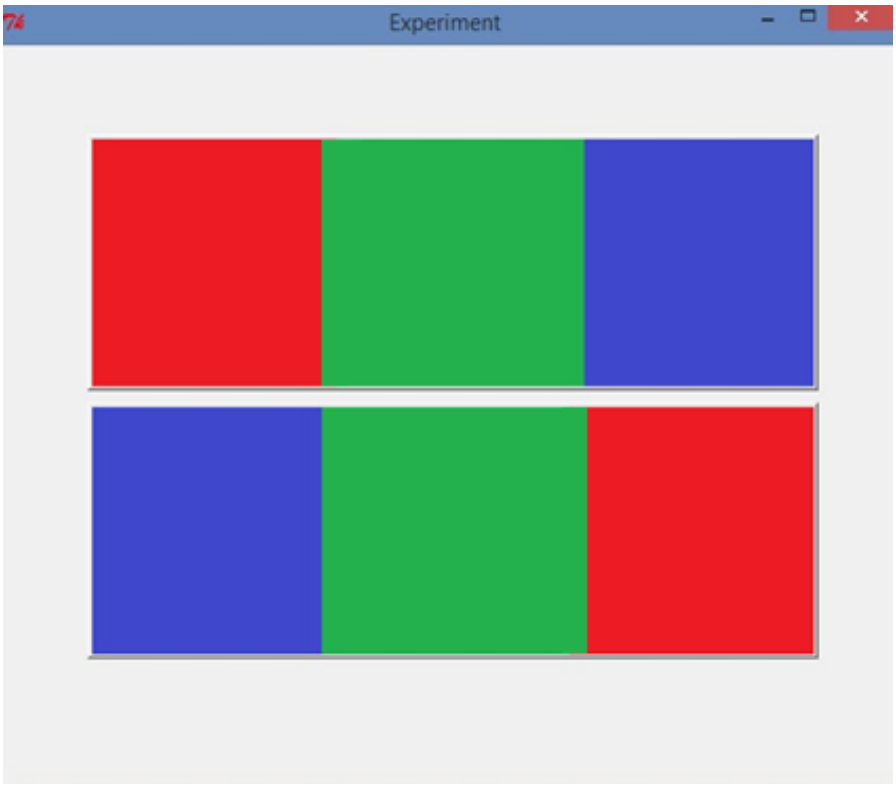

Figure 2.9. Appearance of buttons participants clicked to indicate order of color observed in experiment 2.

Results

The mean color processing speed threshold for processing clockwise versus counter-clockwise direction was statistically the same for all colors at M = 19.15Hz per cycle, 76.60 individual LED flashes per second, F (5, 45) = 0.63, p = 0.68, d = 0.17. Relative to the where processing experiment, the mean color processing threshold for processing hue direction was 2.28Hz per cycle (6.84 individual colors per second), Figures 2.10-2.11. These two processing speed rates differed significantly, $\Delta\mu$ = 16.87Hz, t (9) = 11.31, $p < 0.001$, d = 4.97. A correlation between scores on both tasks was not significant after removing a single outlier who performed exceptionally well on the where processing task, r = -0.086, n = 9, p = 0.826.

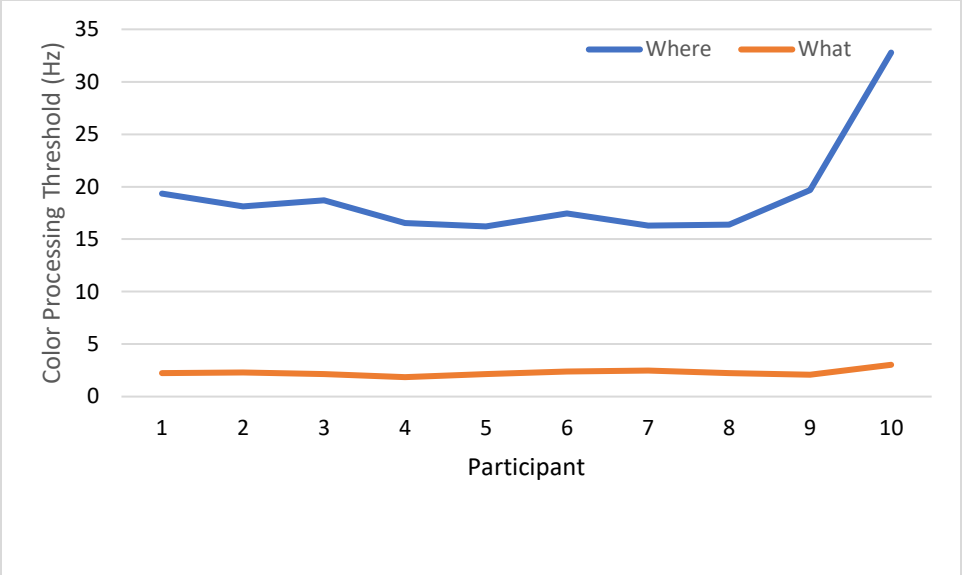

Figure 2.10. Color processing threshold by participant. Those in the *where* experiment show a consistently higher color processing threshold compared to those in the *what* experiment.

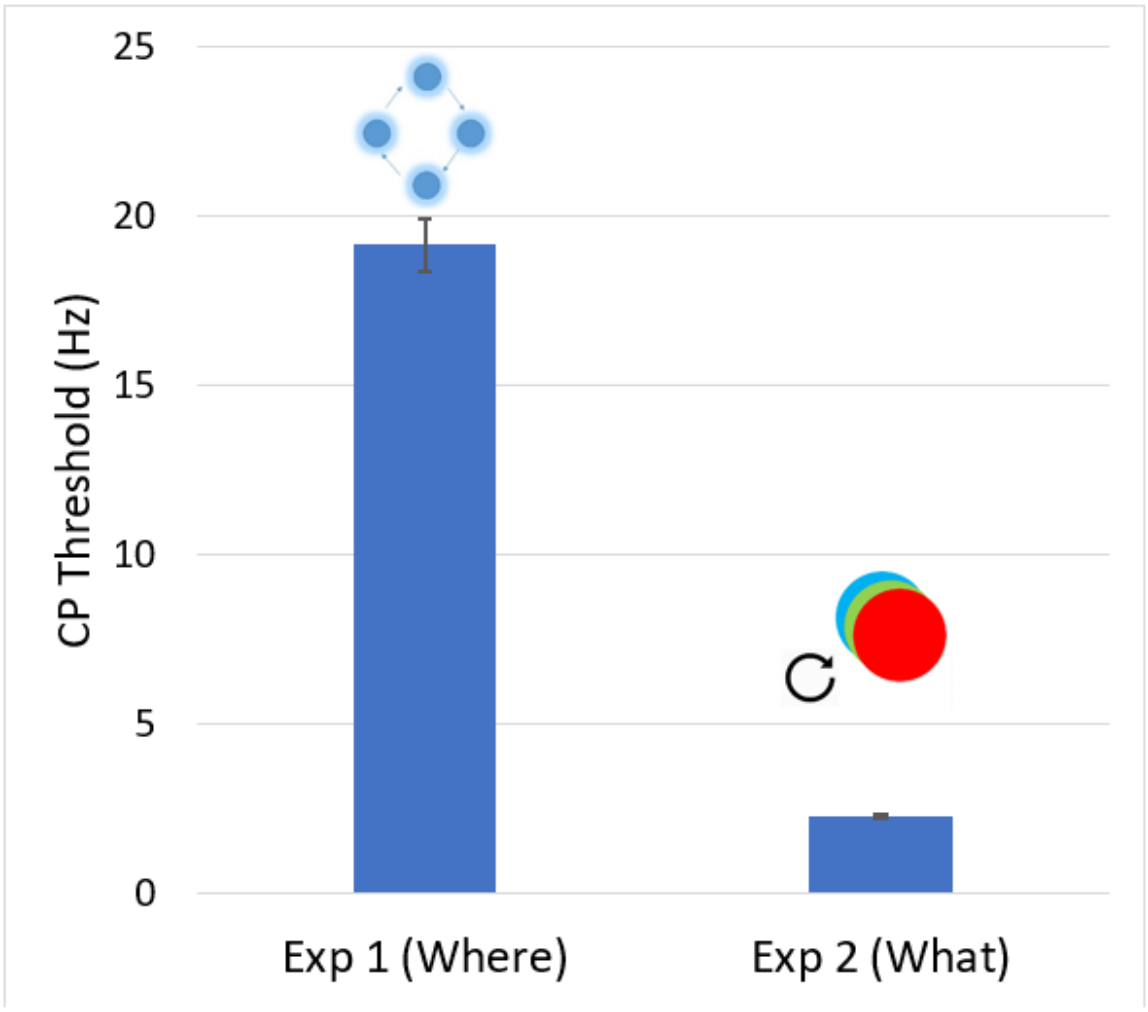

Figure 2.11. Color processing threshold of experiment 1 compared to experiment 2. Because of the increased principally cognitive demand, the threshold decreased in Experiment 2. Standard error bars are shown above each bar.

Discussion

These results demonstrate that there are two distinct timing thresholds for processing temporal color information, supporting my hypothesis that a *what* system processes visual information more slowly than a *where* system. Specifically, the *where* system confirmed processing rate of around 20Hz, consistent with classic findings of flicker fusion rates (Hecht & Shaler, 1936; Simonson & Brozek, 1952) as well as

previous findings on visual temporal rates (Fink et al., 2006). These findings also reveal a *what* system processing rate of about 2-7Hz, consistent with preferred temporal rates of higher-level cognitive processes like shape discrimination judgments, preferred speech syllable rate, and speech processing thresholds (Holloway & McBeath, 2013; Picton, 2013; Van Wassenhove et al., 2004; Van Wassenhove et al., 2007). These results further add to the growing body of literature demonstrating that there is not a unitary temporal processing system as defined by others (Pöppel, 1997). Because of the similar timing rates, these results suggest it is not simply a matter of different levels of difficulty of tasks but in fact supportive of two distinct processing systems. Neuroscientific research also supports the idea of two separate systems, with V1 and V2 processing fast stimuli around 18-25Hz, *where* stimuli, while V4 and the Fusiform Face Area (FFA) process slower temporal stimuli at 4-5Hz, *what* stimuli (McKeeff et al., 2007). Thus, the V1 and V2 areas are likely mechanisms for processing *where* system type stimuli and the V4, FFA are responsible for processing *what* system type stimuli.

A considerable degree of overlap exists between the *where* and *what* visual processing systems and magnocellular (M) and parvocellular (P) visual pathways in the brain (Kessels et al., 1999; Kraus & Nicol, 2005). Even now, creating stimuli that can measure each the M and P systems respectively has proven problematic, especially when there is still much to be discovered in regard to how the two systems contribute to spatial vision (Edwards et al., 2021; Masari et al., 2020). The stimuli used in the current study could be used to more precisely model how the M and P systems function, with these results giving precise processing limits that may apply to these two systems. It should be noted that this work provides a behavioral way of defining these systems from the ground

up, which should lead to cleaner models with more precise thresholds. Future research would determine more precisely what parts of the brain are responsible for behavior related to timing, although current research increasingly points toward temporal processing being a rather dispersed, hierarchical system (Lad et al., 2020).

These findings also mirror past work that shows perceptual-action processing limits of around 20-40Hz for several motor coordination-based tasks (McBeath & Krynen, 2015). Defining perceptual processing limits is important because these factors are necessary in a multitude of crucial perceptual processes such as speech and walking. This dual timing model approach shows promise with this kind of distinction and aligns with other research that demonstrates distinct cortical processing *what* and *where* systems (Creem & Proffitt, 2001; Milner & Goodale, 2008; Rauschecker, 2018). Ultimately, the findings of this study support the differentiation of two visual processing systems and provide a metric for neural research that examines relative dorsal (20Hz) and ventral (2-7Hz) processing rates.

CHAPTER 3

SEEING FALSE COLOR EVEN BETTER: ENHANCED PERCEPTION OF BENHAM'S TOP

First discovered in 1826 by a French Monk by the name of Prévost and then studied scientifically in 1838 by Fechner (Cohen & Gordon, 1949), the idea of seeing "false colors" (FCs) has a centuries-long history. In essence, a black and white patterned disk (Figure 3.1) is spun at an optimal speed to cause the observer to perceive color (Fechner, 1838). Seeing the novelty of the item, Charles Benham marketed the top to the public as a children's toy in 1894, however shortly after its release to the public, it drew the attention of several scientists who attempted to standardize what FCs are commonly seen on the top (Benham, 1894; Bidwell, 1897; Finnegan & Moore, 1894). It is now known that a reddish-brown appears on the outside, followed by tan, then a green and finally a dark blue when spun counterclockwise. The ordering of the FCs reverse position when the disk is spun clockwise, though due to individual differences, not everyone sees all the rings of FCs.

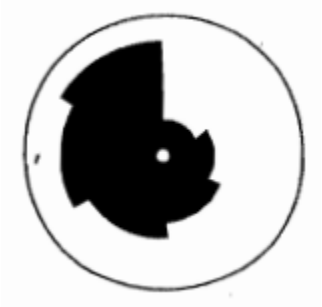

Figure 3.1. Fechner's original disk. This pattern is rarely used in contemporary studies of false color.

Part of why Benham's top tended to be contentious among scientists in the late 1800's and early 1900's—indeed part of the novelty among non-scientists--has to do with how verbal report of the FCs seen on the disk can vary widely (von Campenhausen & Schramme, 1995). However, when colored stimuli are allowed to be matched up to the FCs seen on the disk, individual differences subside with observers typically reporting perceiving the same colors on the spinning disk (Rosenblum et al., 1981).

More modern research points to the idea of the phenomenon being induced by flickering of light between the black and white areas of the disk, and it is now commonly labeled as pattern-induced flicker colors, or PIFCs (von Campenhausen, 1973).

Fechner initially believed the illusion was due to timing with red being perceptually processed more quickly and blue more slowly; hence when viewing a disk of varying black and white patterns, rapid changes in luminance would create the illusion of PIFCs (Fechner, 1838; von Campenhause & Schramme, 1995). Variations of this theory persisted with some attributing the effect of PIFCs to how cones are placed in the retina (Jarvis, 1977; Kellog & O'Shea, 2007). More recent work shows that the effect is better explained as the result of high order visual processing of the same rapid changes in luminance (Le Rohellec & Vienot, 2000). Using fMRI, it has been demonstrated that seeing PIFCs are largely processed in the early visual areas V1 and V2, and some further processing in V4 (Mullen et al., 2015; Tanabe et al., 2011). The same work with Benham's tops demonstrates that both color and contrast processing are linked (Kenyon et al., 2004; Mullen et al., 2015). This idea warrants further exploration using graded top patterns to psychophysically test for whether contrast will affect perceived FCs.

Further complicating the phenomena, the ambient lighting spectrum has been shown to influence perceived FCs on a Benham's top (Vienot & Le Rohellec, 1992). Several different light sources were usually originally used when Benham's top first became popular, primarily incandescent bulbs, but Benham himself encouraged others to use a low-pressure sodium lamp to observe the top to most successfully see PIFCs (Benham, 1894). Yet, this idea has not previously been formally tested, especially with the more modern knowledge that low pressure sodium lamps put out an extremely narrow frequency band of light (de Groot & Van Vliet, 1986).

Several new designs of the top have emerged over the centuries, and the design of the disk itself has been inconsistent between studies with some patterns having seen more use than others, likely due to evoking PIFCs better than others (Nishiyama, 2012). It is potentially useful to more precisely test the range of PIFCs under conditions of different tops and lighting, and to directly test if the more commonly used top patterns objectively create better false colors then the less common designs. Another possible fruitful area to explore is how these dimensions map out using colorblind observers. There is reason to believe that colorblind observers will have suppressed sensitivity to false color based on previous work (Rosenblum et al., 1981;Shepard & Cooper, 1992; White et al., 1977), but it is possible that even still, perception of false color can be improved by altering the nature of the stimulus and surroundings.

The current experiment seeks to test if top pattern type as well as narrowband ambient lighting can enhance the FCs seen on a Benham's top, both in terms of number of unique FCs seen on the disk as well as vividness of the FCs. Further, the uniformity of the stimulus pattern itself has not been tested. This leads to four hypotheses:

H1: Increased clarity of the top pattern will produce both a broader range of perceived FCs and higher saturation of FCs.

H2: Narrowband ambient lighting will enhance the number of FCs and saturation of FCs.

H3: More commonly used top patterns will produce more vivid and saturated FCs.

H4: Colorblind individuals will experience fewer vivid FCs than normal trichromats overall.

Method

Participants

4 participants were recruited for the experiment, 3 female and 1 male, (M_{age} = 20.25, *SD* = 0.50). More participants were planned, but testing limitations occurred due to Covid-19 restrictions. All participants had normal or corrected-to-normal vision and were trichromats.

An additional 2 colorblind individuals were recruited for the experiment, (M_{age} = 19.50, *SD* = 0.71), one deuteranope and one protanope.

Materials

A 3.5” diameter (8.9cm) custom formed plastic disc was installed into a photo studio light box (Figure 3.2) with viewing hole and connected to a 12v DC stepper motor spinning counterclockwise at 300rpm with rotational speed being verified by a built-in tachometer using a Hall effect sensor (8 micron [.008mm] wobble); this speed was chosen due to creating the best FCs (von Campenhausen, 1973). To control speed, a

variable output power supply was attached to the motor and by varying voltage, we could maintain precise speeds of the top. The top was spun counter-clockwise on all trials, since changing directions also changes the location of FCs such that the FCs reverse locations, that is, the outer FC will appear on the inner portion of the top and vice versa with the inner FCs appearing on the outer portions of the top.

Patterns for the top were printed onto the HP Premium 32 paper due to the excellent contrast provided by the 100 brightness of the paper (Figure 3.3). Three patterns were chosen, the original Benham's Top pattern (1894), a popular pattern created by von Campenhausen (1995), both of which are known to work splendidly, and a top pattern that does not typically work so well, but is known to produce some false color based on pilot testing. To test for the possible confound of reflectivity, the reflectance curves for each the black and white areas of the top were measured using an Ocean Optics USB 2000 spectrometer, (Figure 3.4). Both the black and white portions remain relatively uniform across all visible ranges of the spectrum with the white portion having a small peak in the 435nm range of the spectrum with a drop-off at lower wavelengths past that same mark, a common trait of white paper.

Patterns were shown with sharp edges or a 25% blur using Photoshop 2019 (Figure 3.5). From the roof of the box, an ASU SciHub HyLighter device was mounted with its light output reflecting off the disc at a 45º angle (Krynen & McBeath, 2021). Six colored narrowband monochromatic LEDs were chosen to eliminate the potential confound that seeing PIFCs could be a function of chromatic ambience, rather than simply as a function of narrowband ambient illumination. Each LED was tuned to reflect exactly 22 Lux from the surface of the top following prior research with Benham's tops

(Vienot & Le Rohellec, 1992) and also because varying luminance levels have been shown to affect what FCs observers may see (Chen & Cicerone, 2002). A chin rest was mounted 6in (15cm) in front of the window of the box (23.5in, 63.5cm from the spinning top itself) to keep viewing distance constant and stimulate the paracentral area of the fovea, 8° field of view.

To assist in determining number of FCs seen, a bridge was placed above the spinning disk with equally-spaced 10mm tick marks and numbers so participants could record approximately where a FC ring was located, Figure 3.6. Number of unique hues in each trial were then calculated, with the bridge being split into 4 quadrants to reflect each of the four arcs featured in the patterns chosen, Figure 3.7. This was particularly important as some of the top patterns produced thicker observed FCs and hence only one FC per quadrant was used for final analysis. Quadrants consisted of first, area 1 (the outermost quadrant) containing marks 2-3, followed by area 2, marks 4-5, then area 3, marks 6-7, and finally area 4 (the innermost quadrant), of marks 8-9. No false color was observed at the outermost edge of the disk or the innermost centroid of the disk and hence the marks 1 and 10 are excluded from the present analysis.

Next, to provide a quantitative measure for saturation of hue perceived, Color-Aid Corp colored paper slips were cut into 1x1" squares and affixed to a board next to and outside of the spinning top box then labeled (for ease of participant choosing colors) in six "chunks" of the prototypical colors identified from experiment 1: red, orange, yellow, green, blue and violet (Appendix B). Within each color chunk, lightness and saturation were represented on the x- and y-axes respectively for a total of 54 colors with white, black and four increasingly dark shades of gray meaning participants had 60 possible

options from which to choose. The Color-Aid Corp panels were illuminated to 150lux by a 5000k broadband Waveform Lighting Absolute Series LED light strip to closely approximate horizon daylight, Figure 3.8. An additional benefit is the Color-Aid Corp values of the color swatches can be converted to other standardized color measures such as the Natural Color System (NCS) or CIE color space (Arbab et al., 2018). Each of the possible color options is shown in a table with corresponding CIE 1931 *x, y* values and plotted in the CIE 1931 color space in the appendix. In order to quantify saturation of perceived false color, first the white point was determined using the CIE coordinates of the "WHITE" Color-Aid Corp paper ($x = 0.3187$, $y = 0.329$). Distance from white then represents saturation with more fully saturated colors residing at the edges of the color space (Choudhury, 2014; Fortner & Meyer, 1998). Therefore, cartesian distance, *d*, between each of the color options and the white point were calculated for each trial using the following formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

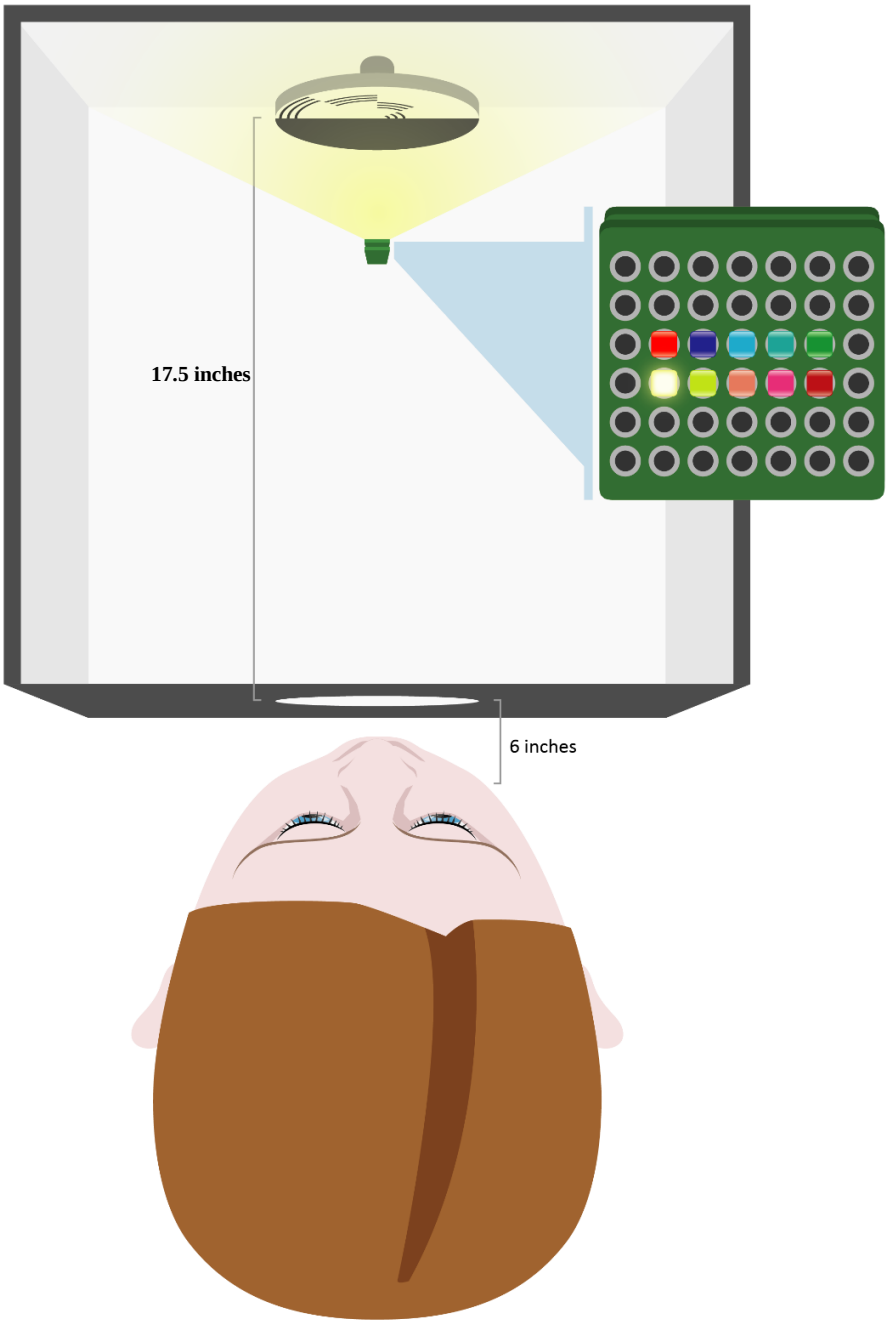

Figure 3.2. Top-down view of apparatus created to test Benham's top patterns. The observer looks through a window and sees the spinning disc with a light attached to the roof of the box shining down at a 45º angle off the surface of the spinning top. View of LED panel is rotated in diagram to show arrangement of LEDs. In figure, the yellow LED is illuminated.

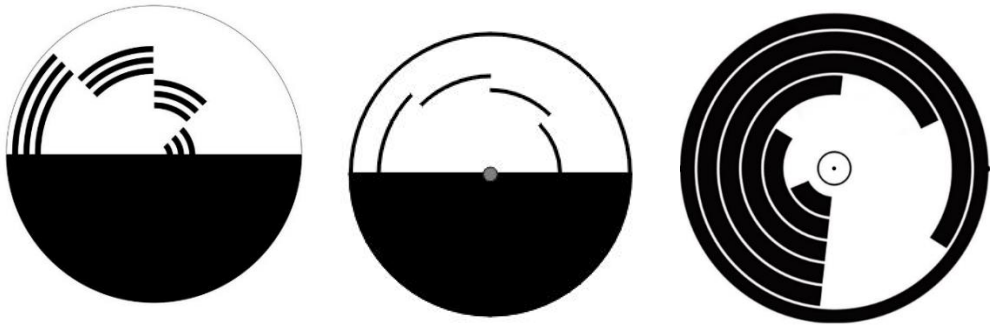

Figure 3.3. Three top patterns that were used in experiment 4. The left and center patterns have demonstrated the most variety of false color as well as the greatest saturation by pilot testing whereas the right pattern was less successful in producing false colors.

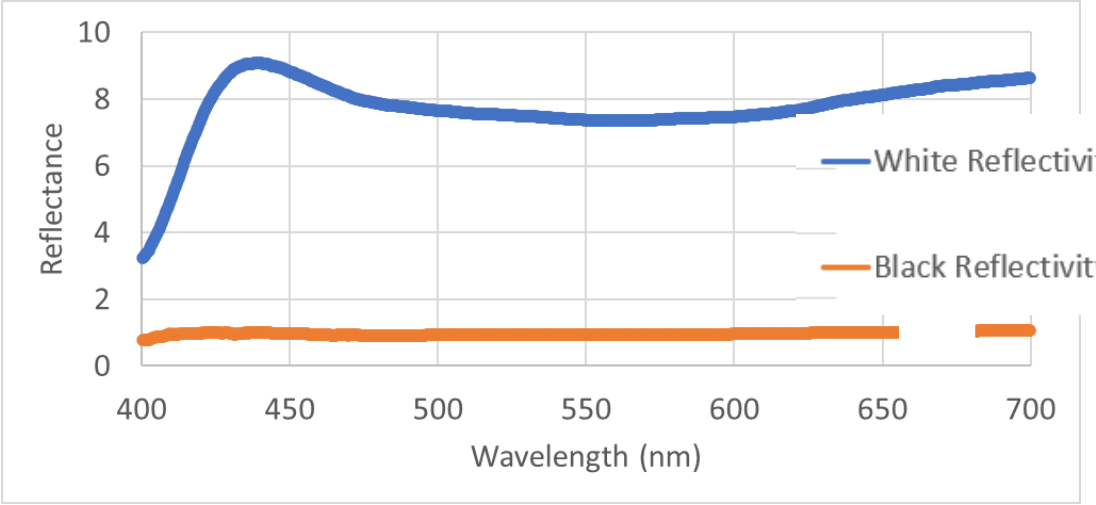

Figure 3.4. Reflectance curves for the white and black portions of the top. Both the black (lower curve) and white (upper curve) remain relatively uniform across all visible ranges of the spectrum with the white portion having a small peak in the 435nm range of the spectrum with a drop-off at lower wavelengths past that same mark, a common trait of white paper.

Figure 3.5. Blurred top patterns. Patterns were shown to participants at 25% blur.

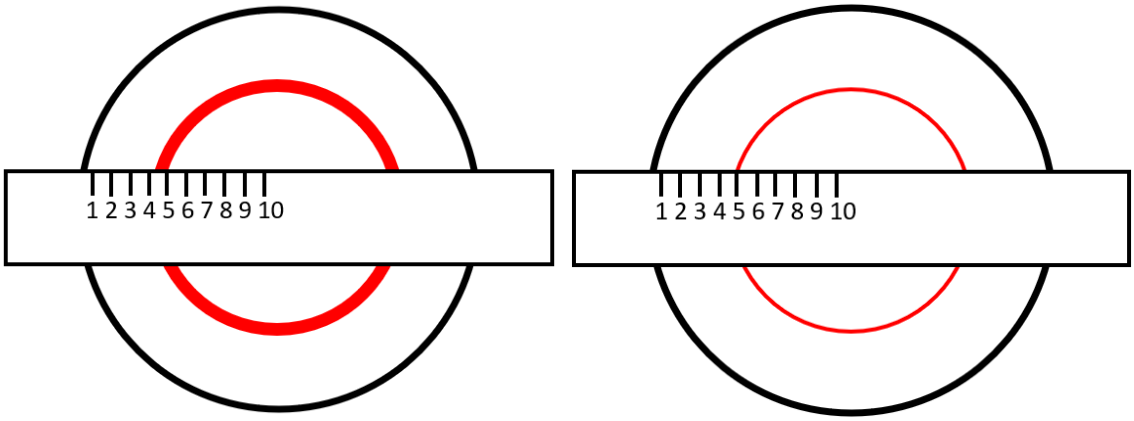

Figure 3.6. Diagram of bridge placed over top with numbered tick marks. This enabled participants to precisely identify where false colors were seen, in this example the color red would be reported in the #5 location. The top on the left represents how the original Benham's top typically produces a thicker FC than the top on the right, used by von Campenhausen.

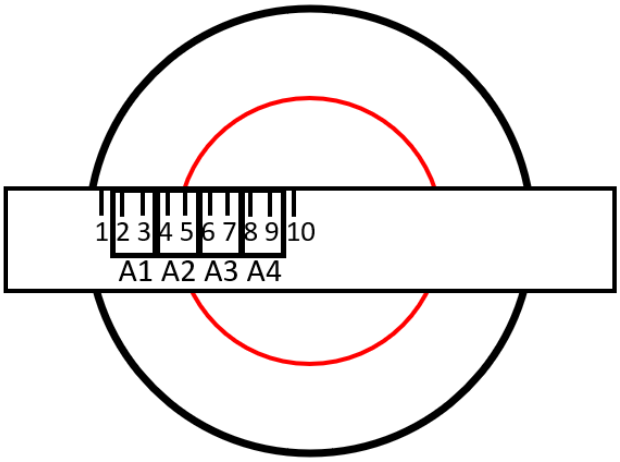

Figure 3.7. Depiction of how quadrants were defined for analysis. Quadrants were used to account for some lines being thicker based on pattern type as well as to quantify what false colors were seen on each of the arcs.

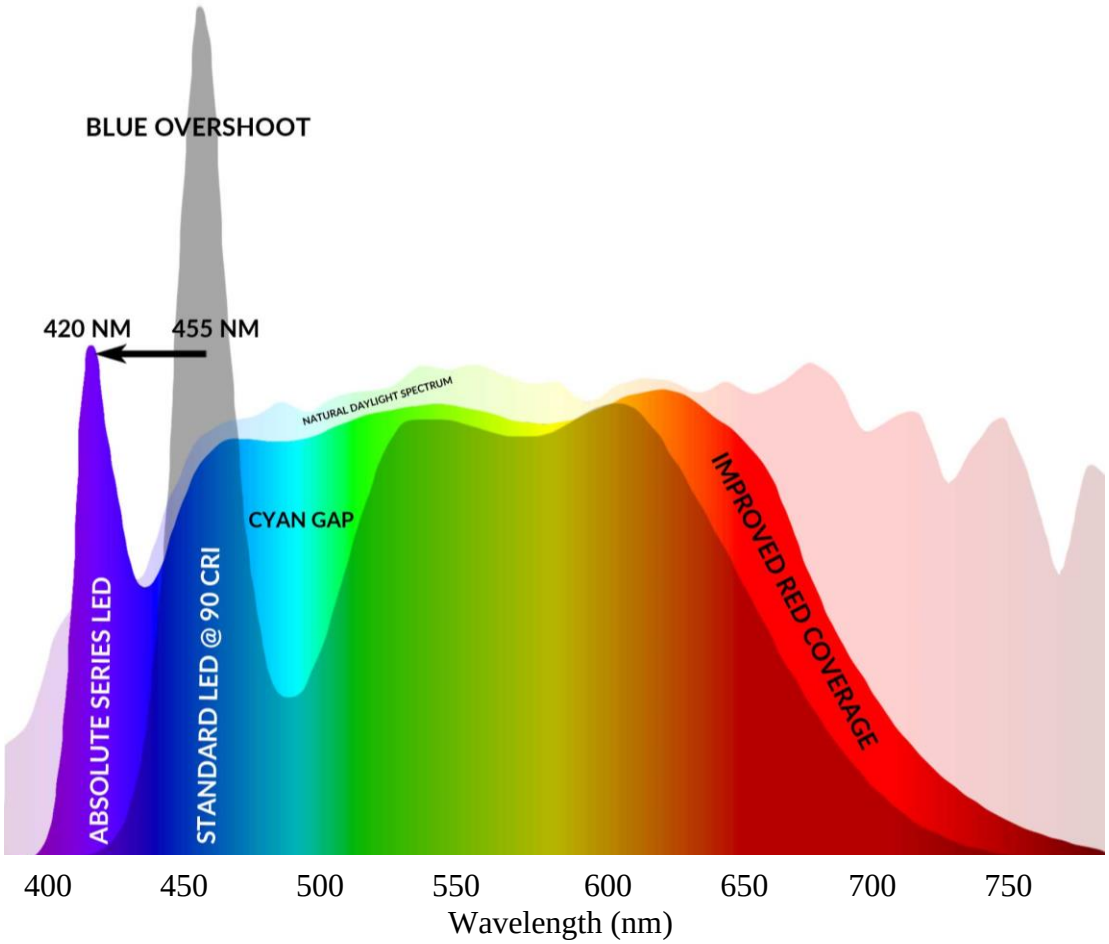

Figure 3.8. Lighting spectrum of broadband 5000k light source. This source was chosen for its excellent coverage of the natural daylight spectrum, with improved red coverage as well as better coverage of the 470-530nm zone (commonly called the cyan gap) when compared to traditional broadband LEDs.

Procedure

Participants sat in front of the box and observed the spinning top through the viewing hole. Using a nearby keyboard and computer screen, participants compared the FCs on the spinning top to the panel of Color-Aid Corp colored papers then reported which hues were seen for each ring in a pre-made electronic spreadsheet (Figure 3.9). To begin, participants would randomly see one of the three possible disk patterns illuminated

with 5000k broad spectrum light to maximally elucidate false colors, then participants would see the disk illuminated in the following order: Red (660nm), Orange (613nm), Yellow (587nm), Green (530nm), Blue (473nm), then Violet (450nm), followed by a disk change to the 25% blurred pattern of the same disk. This process was repeated for the other two disk patterns presented in random order until all trials were completed, taking about an hour in total. An example trial would consist of a participant placing their head on the chinrest and observing the spinning disk, then comparing the FCs on the disk to the Color-Aid Corp squares, and inputting into a pre-made spreadsheet what labeled color square most closely matched the FC perceived at each labeled mark on the disk. No time limit was placed on when participants were to respond for perceived FC, with each trial taking on average 30 seconds to one minute to complete. After all possible illuminations were presented in the above-mentioned order, the disc pattern was switched out to its blurred counterpart until all possible illuminations were shown on the blurred pattern. A new pattern would be randomly selected and the experiment would proceed until all trials were presented and recorded by the participant.

The participants viewed 2 edge gradations, with 7 lighting conditions and 3 pattern types for a 3 x 7 x 2 design.

Figure 3.9. Photograph of experimental setup. The participant is inputting data into the computer after viewing the top.

Results

Number of unique FCs and saturation of FCs were assessed to test for how well disc clarity, ambient lighting and pattern type contributed to the Benham's top illusion.

For the normal vision participants, the mean number of FCs seen in the clear top conditions was 2.53 (*SD* = 0.51) and in the blurred conditions was 1.88 (*SD* = 0.69). Overall, participants saw the least number of FCs in the natural lighting condition, Figure 3.10. Using a three-way ANOVA examining the effect of clarity, lighting, and top on number of perceived FCs, there were two significant simple main effects of Clarity, F (1, 126) = 19.77, p = 0.001, d = 0.87 (Figure 3.11) and Lighting, F (6, 126) = 9.27, p = 0.001, d = 1.77 (Figure 3.12). A Tukey's post hoc test indicated that number of FCs seen

in the natural lighting condition was significantly fewer than each pairing of individual narrowband lighting colors: red, orange, yellow, green, blue, and purple ($p < .001$).

However, the same three-way ANOVA demonstrated there was no significant three-way interaction between clarity, lighting and top on number of FCs, F (12, 126) = .48, p = 0.92, d = 0.41 or two-way interactions, Lighting x Clarity, F (6, 126) = .41, p = 0.86, d = 0.29; Lighting x Top, F (12, 126) = 0.64, p = 0.81, d = 0.51; and finally, Clarity x Top, F (2, 126) = 0.18, p = 0.80, d = 0.09).

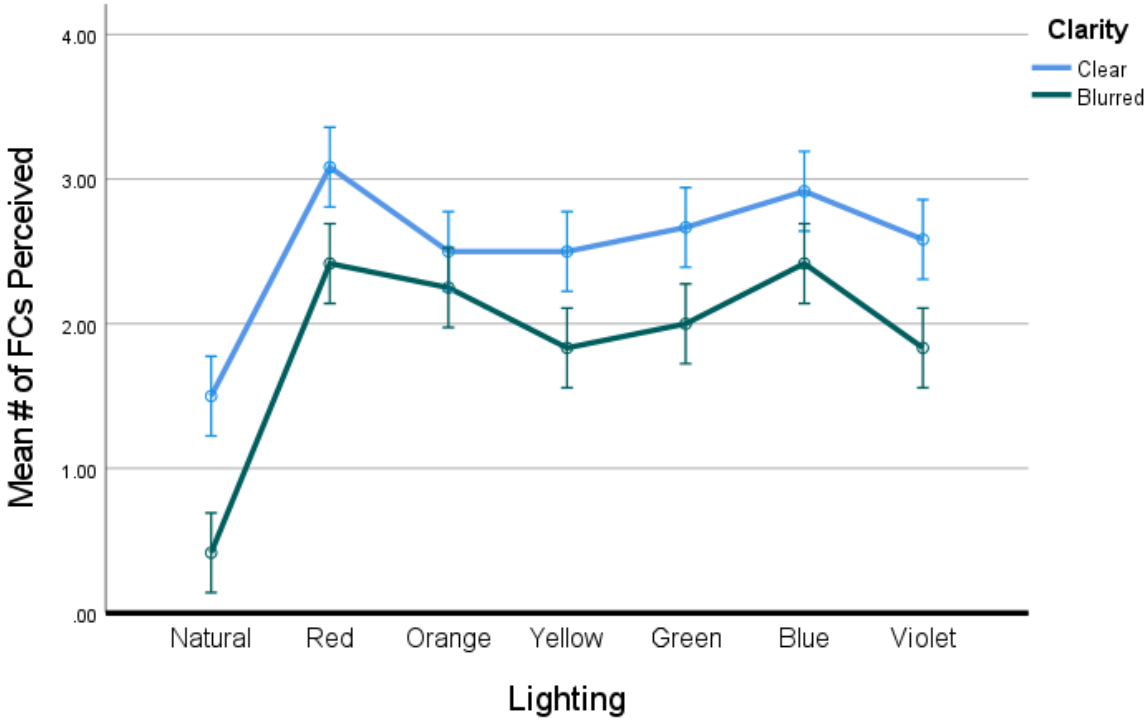

Figure 3.10. Average number of FCs observed by condition and clarity of top pattern. Using narrow-band light enhanced the number of FCs participants observed.

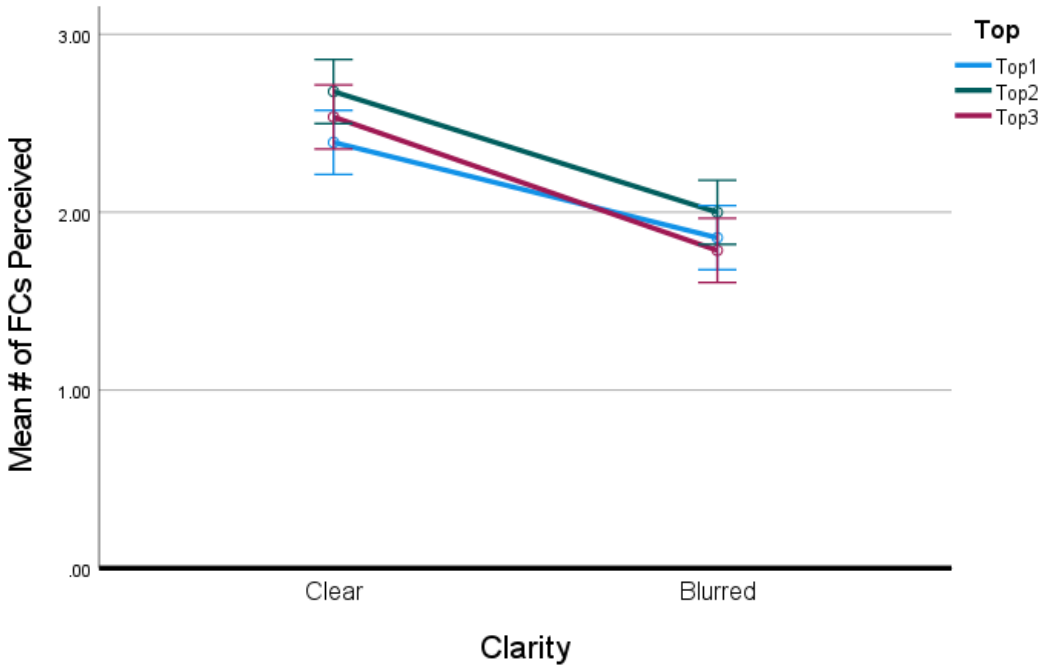

Figure 3.11. Mean number of FCs perceived by top type and clarity. There was a significant effect of clarity on number of FCs seen.

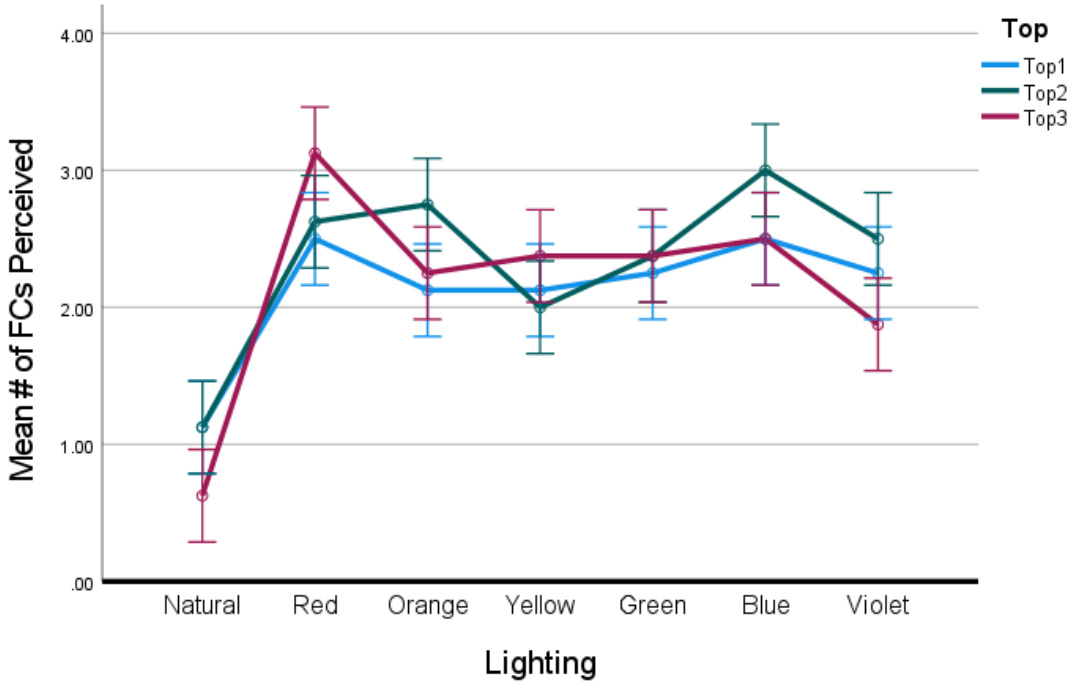

Figure 3.12. Mean number of perceived FCs between the three top patterns and ambient lighting conditions. No significant effects of top pattern were observed in number of indicated FCs.

Saturation values ranged from 0.0003 to 0.2792 from white with a mean saturation value of 0.12 (*SD* = 0.02) in the clear top conditions and a mean saturation value of 0.10 (*SD* = 0.03) in the blurred top conditions (Figure 3.13). Participants saw the highest saturated false colors in the narrowband colored wavelength lighting conditions overall, with pattern #1 exhibiting the most saturated FCs in natural light and the higher wavelengths but then becoming the least saturated pattern toward the lower wavelengths, Figure 3.14.

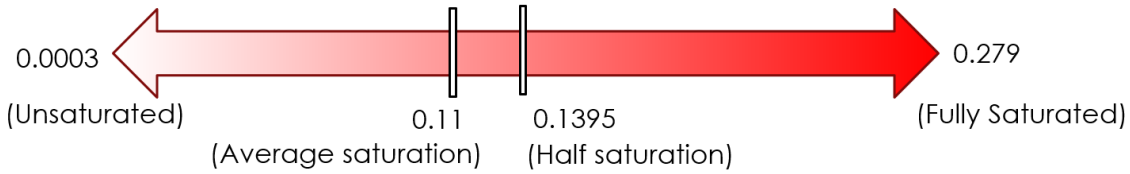

Figure 3.13. Visual illustration for interpreting saturation values. The saturation values ranged from 0.0003 to 0.279, with a possible halfway saturation value of 0.139, note that the actual average saturation value was 0.11, indicating overall less saturated values.

A three-way ANOVA between clarity, lighting, and top based on saturation of false color indicated a significant two-way interaction between lighting and top, F (12, 126) = 2.35, p = 0.009, d = 1.06 with significant main effects of Clarity, F (1, 126) = 8.11, p = 0.001, d = 0.51 (Figure 3.15) and Lighting, F (6, 126) = 12.04, p = 0.001, d = 2.30, Figure 3.16. A Tukey's post hoc test revealed that the mean observed saturation levels in the natural lighting condition were significantly less than each pairing of the narrowband colors: red, orange, yellow, green, blue, and purple (p < .005).

No significant three-way interaction was found, F (12, 126) = 0.87, p = 0.57, d = 0.59 as well as two-way interactions between Lighting x Clarity, F (6, 126) = 0.35, p = 0.91, d = 0.29 and Clarity x Top F (12, 126) = 0.94, p = 0.39, d = 0.29 respectively.

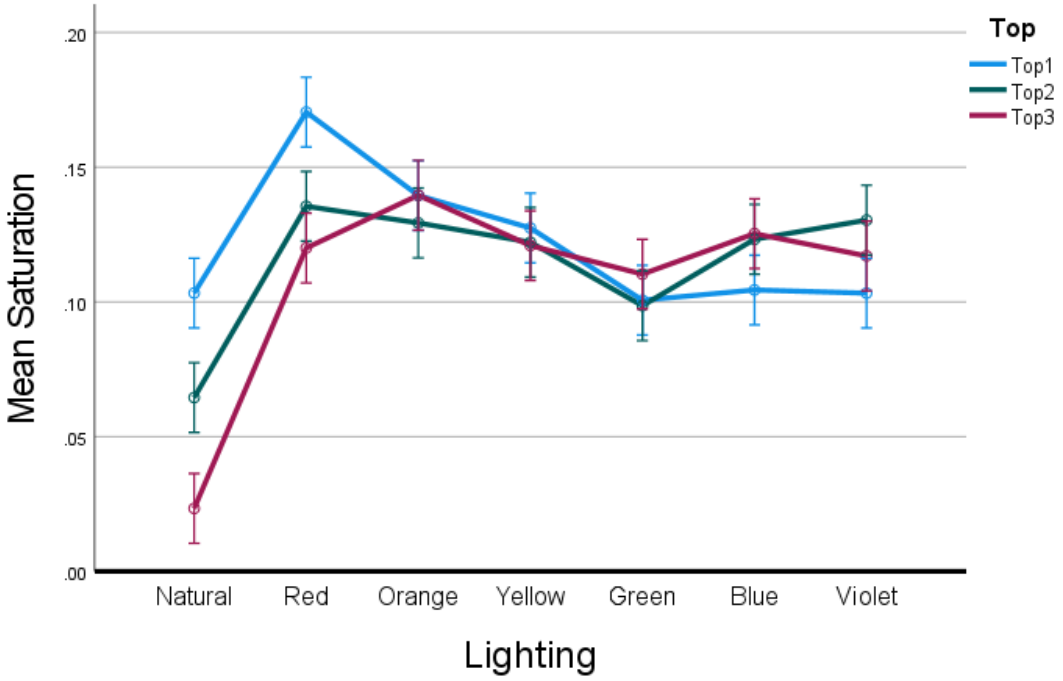

Figure 3.14. Mean saturation values between the three top patterns and ambient lighting conditions. Participants compared the false color from the disc to Color-Aid Corp paper squares which were then converted into CIE color space values.

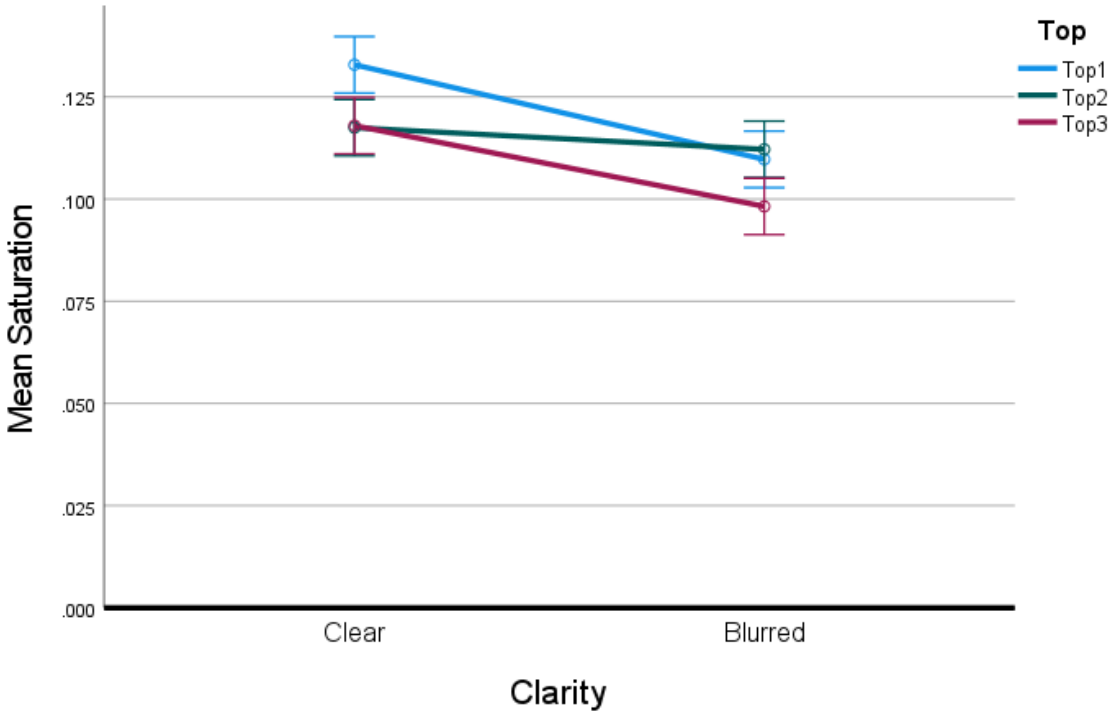

Figure 3.15. Clarity by top and pattern type. As with number of FCs, saturation was also highest when the pattern had a high level of clarity.

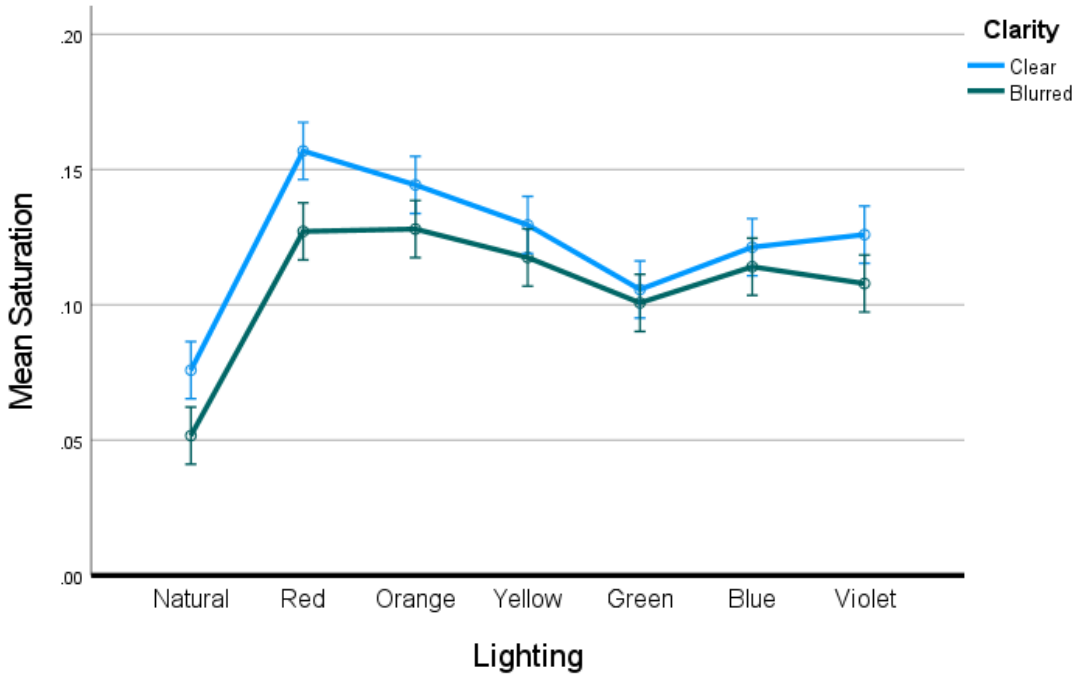

Figure 3.16. Mean saturation by clarity of pattern and color. Narrowband colors induced the best perception of saturated false colors.

The frequency of FCs (red, orange, yellow, green, blue and violet) seen in each quadrant was summed. Red was most commonly seen in the outermost quadrant of the disk (F (5, 71) = 69.23, p = 0.001, d = 3.54) being significantly more observed than any other false color using a Tukey's post hoc test ($p < 0.05$), Figure 3.17. Orange was most frequently observed in quadrant two, F (5, 71) = 33.54, p = 0.001, d = 2.47, also being seen significantly more than the other false colors in post hoc comparison ($p < 0.05$). Green was seen most in quadrant three, F (5, 71) = 51.11, p = 0.001, d = 3.05, and also seen more than others using post hoc analysis. Finally, in the innermost quadrant four, blue was most seen, F (5, 71) = 48.76, p = 0.001, d = 2.97, also confirmed with a post

hoc analysis. The use of quadrants was further justified as no participant perceived more than 4 FCs in any of the top patterns without quadrant constraints.

Overall, red, orange, green and blue were seen more frequently than yellow and violet, Figure 3.18, Appendix B.

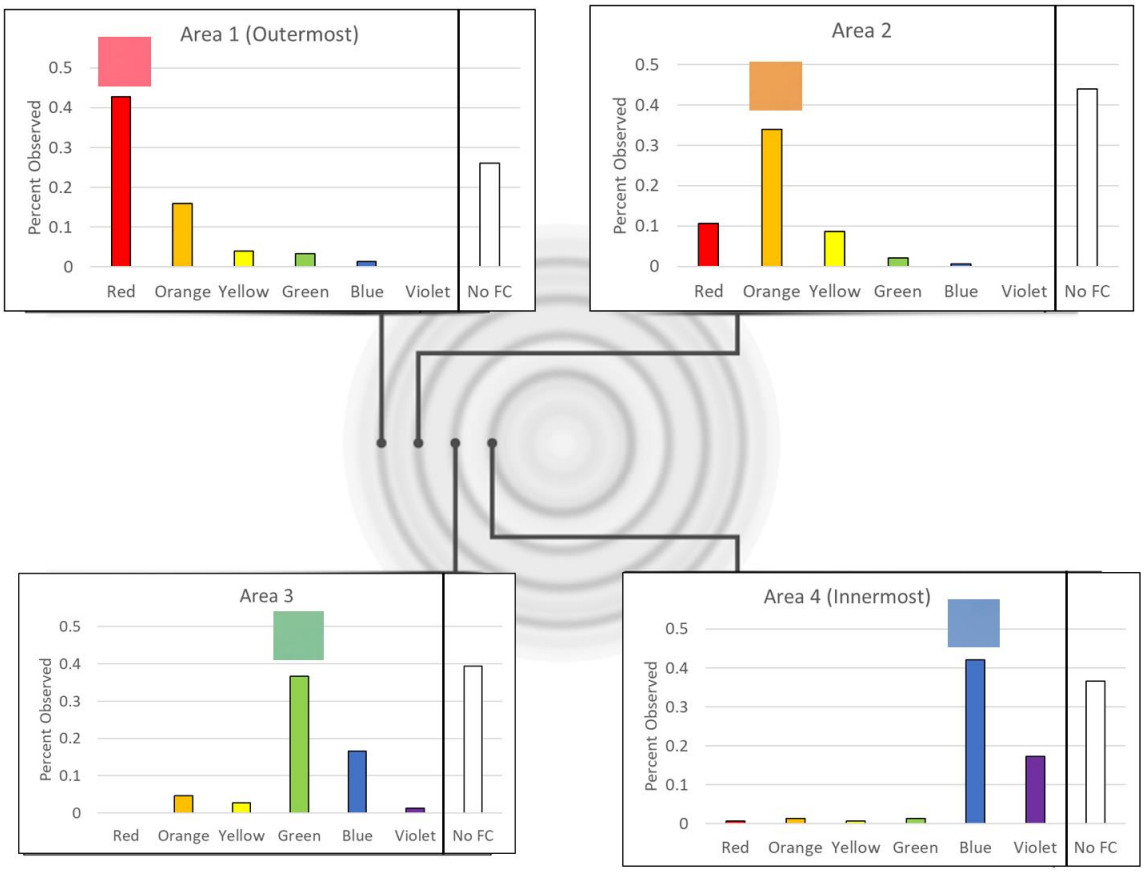

Figure 3.17. Frequency of FCs observed on disk by quadrant by normal vision observers. FCs are separated into quadrants. The number of observations of FCs per area are as follows: Area 1(n = 101), Area 2(n = 84), Area 3(n = 91), Area 4(n = 95). Prototypically seen FCs for each area are featured in each chart.

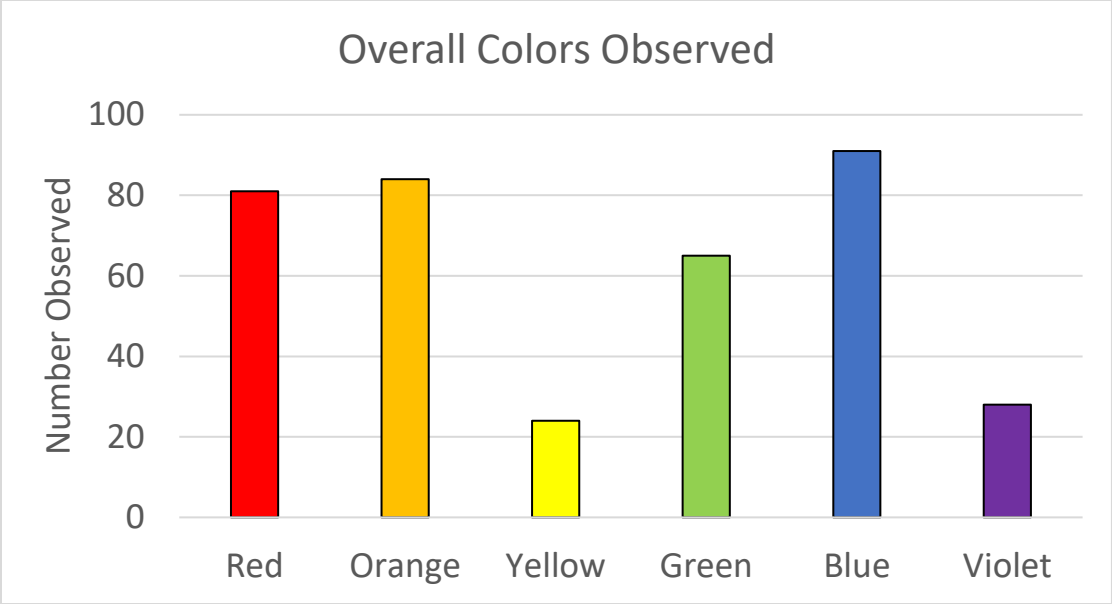

Figure 3.18. Frequency of FCs observed overall by normal vision participants. Yellow and violet were the least observed FCs.

Analysis of colorblind participants revealed that they saw significantly fewer FCs than normal trichromats, M_{FCs} = 1.59, SD = 1.14, t (250) = 4.13, p = 0.001, d = 0.49. A three-way ANOVA was carried out between lighting, clarity, and top was conducted based on number of FCs, showing significant simple main effects of Clarity, F (1, 63) = 18.70, p = 0.001, d = 1.08, and Lighting, F (6, 63) = 16.85, p = 0.001, d = 1.02, Figures 3.19-3.21. In post hoc comparison, unlike normal vision individuals, the number of FCs seen under red narrowband illumination was significantly decreased from all the other lighting conditions ($p < 0.05$) except for the natural light pairing. No interactions or other main effects were found in the three-way ANOVA of number of FCs seen by colorblind participants.

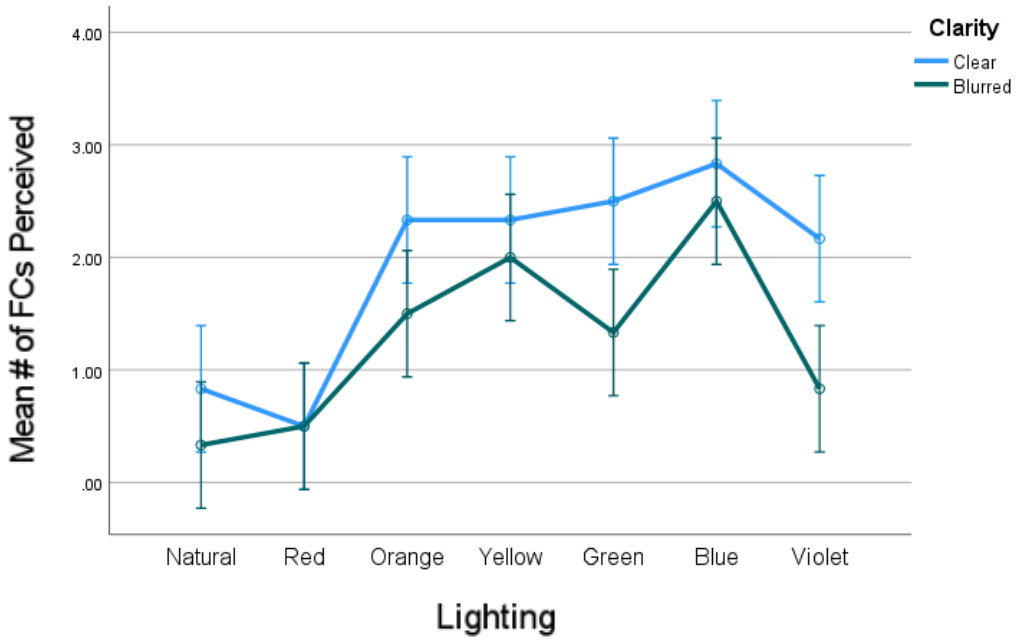

Figure 3.19. Average number of FCs observed by condition and clarity of top pattern for colorblind observers. Clarity enhanced number of FCs seen across lighting conditions.

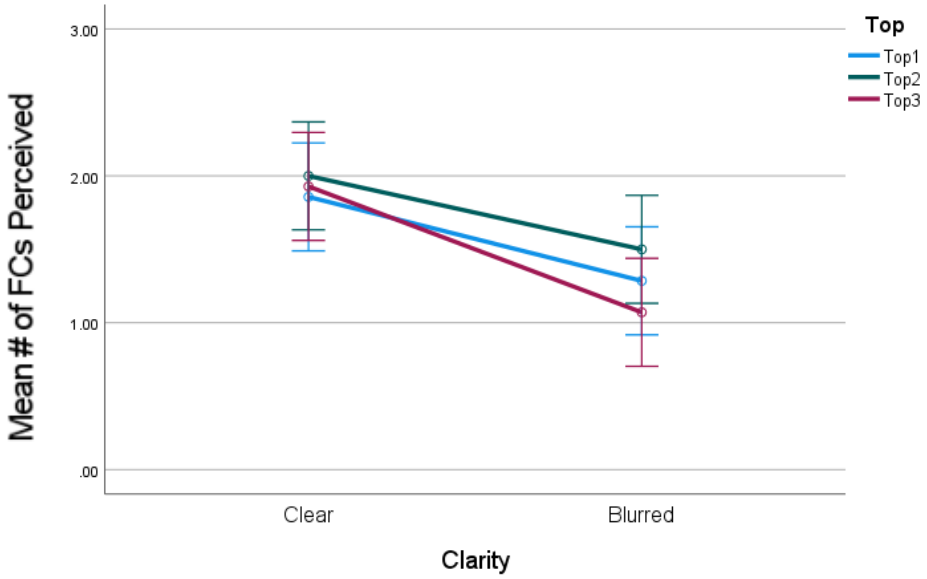

Figure 3.20. Mean number of FCs perceived by top type and clarity for colorblind observers. Clarity enhanced number of FCs seen across top patterns.

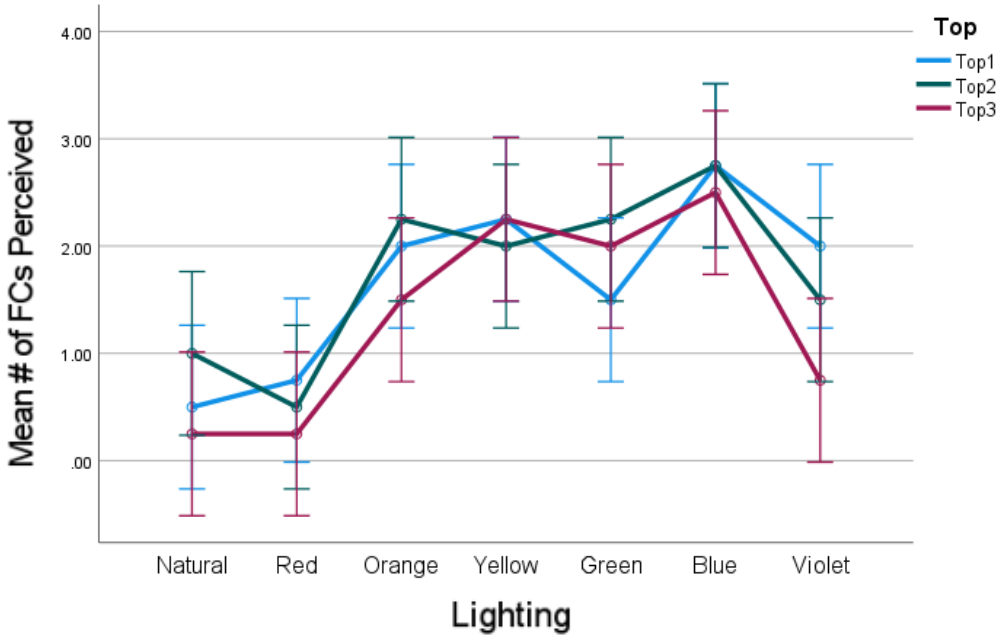

Figure 3.21. Mean number of perceived FCs between the three top patterns and ambient lighting conditions for colorblind observers. Natural and red narrowband lighting had similar effects on number of FCs perceived.

Saturation scores of colorblind participants were also decreased significantly compared to their normal vision counterparts, t (250) = 7.59, p = 0.001, d = 1.01. A three-way ANOVA on the effects of lighting, clarity and top as a result of saturation, showed two significant simple main effects of Clarity, F (1, 63) = 5.44, p = 0.025, d = 0.58, and Lighting, F (1, 63) = 20.93, p = 0.001, d = 1.14, Figures 3.22-3.24. Post hoc analyses revealed the natural and red lighting, again, were significantly different from the other ambient lighting conditions ($p < 0.05$) but not from each other. Blue and orange ambient lighting also had a significant difference from each other in saturation of FCs perceived with FCs in the blue condition appearing more saturated (p = 0.01). No interactions or other simple main effects were present in the three-way ANOVA on saturation seen by colorblind participants.

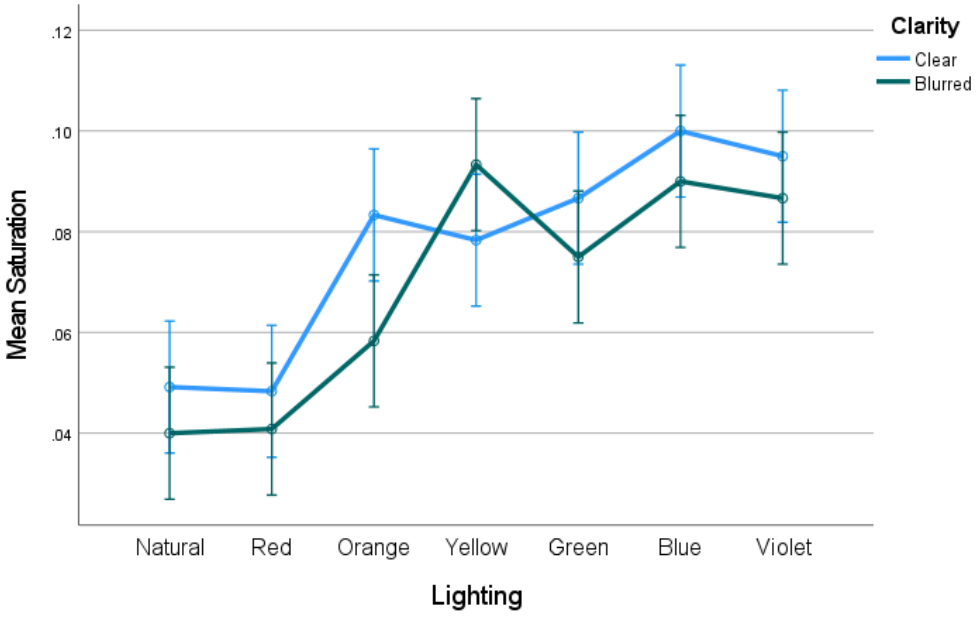

Figure 3.22. Mean saturation by clarity of pattern and false color for colorblind observers. Natural and narrowband red lighting affected saturation of viewed color.

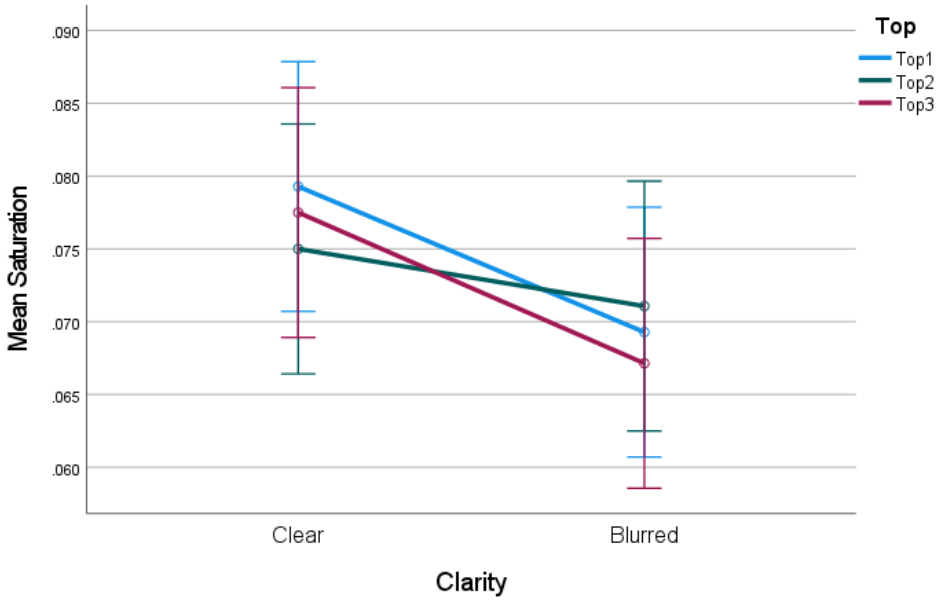

Figure 3.23. Clarity by top and pattern type for colorblind observers. The graded patterns caused colorblind participants to perceive less saturated FCs.

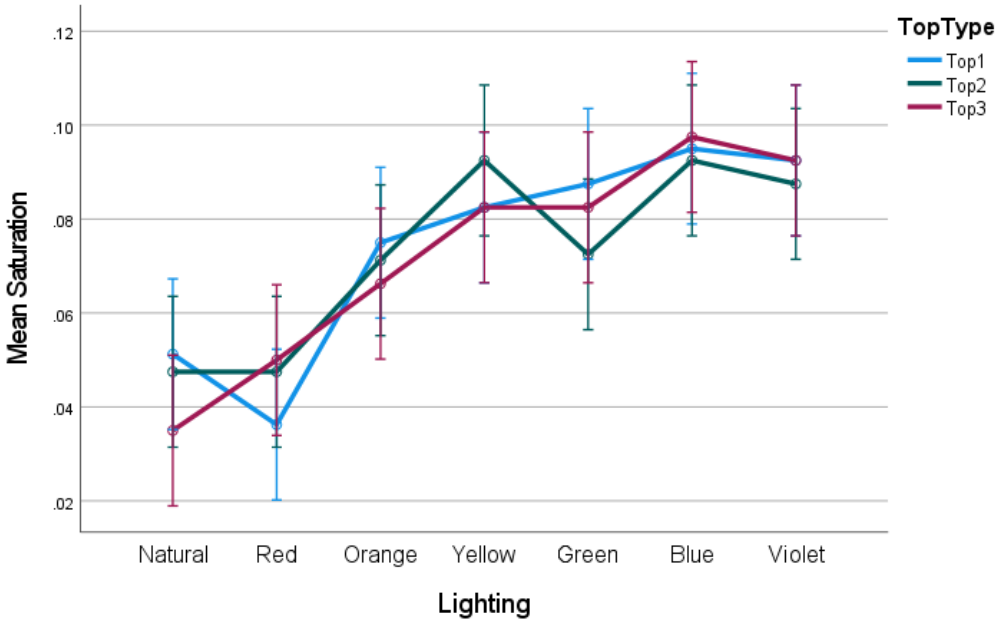

Figure 3.24. Mean saturation values between the three top patterns and ambient lighting conditions for colorblind observers. Top pattern was not significant in perceived saturation.

Colorblind observers' frequencies of FCs (red, orange, yellow, green, blue and violet) seen in each quadrant was summed. Orange was most commonly seen in the outermost quadrant of the disk (F (5, 35) = 10.62, p = 0.001, d = 0.88) being significantly more observed than any other false color using a Tukey's post hoc test ($p < 0.05$), Figure 3.25. Orange was also most frequently observed in quadrant two, F (5, 35) = 16.26, p = 0.001, d = 1.09, also being seen significantly more than the other false colors in post hoc comparison ($p < 0.05$). Green and blue were seen most in quadrant three, F (5, 35) = 6.23, p = 0.001, d = 0.67, and also seen more than other FCs except for from each other and orange using post hoc analysis ($p < 0.05$). Finally, in the innermost quadrant four, blue and violet were seen most frequently, F (5, 35) = 13.15, p = 0.001, d = 0.98, also

confirmed with a post hoc analysis more than the other colors but not from each other ($p < 0.05$).

Overall, orange and blue were seen more frequently than the other possible FCs for colorblind participants, Figure 3.26.

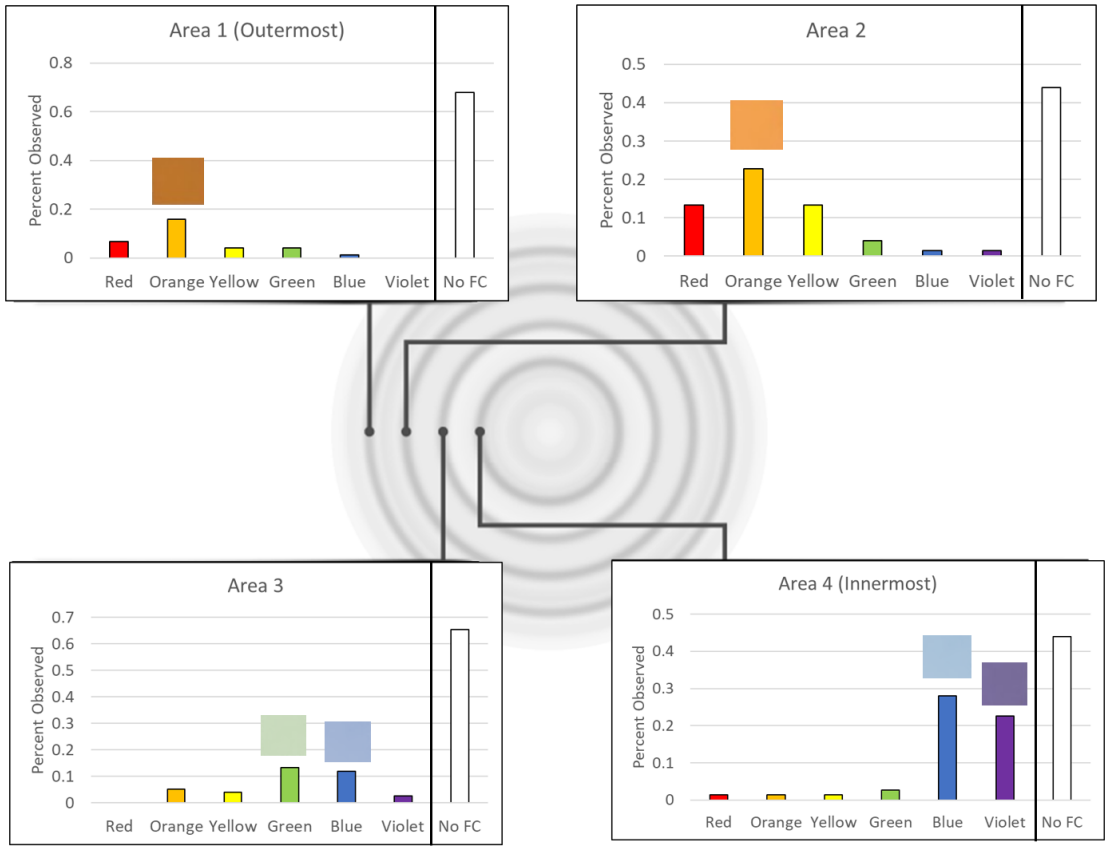

Figure 3.25. Frequency of FCs seen by colorblind observers. The number of observations of FCs per quadrant area are as follows: Area 1(n = 24), Area 2(n = 42), Area 3(n = 26), Area 4(n = 41). Protypically observed colors are featured in each chart.

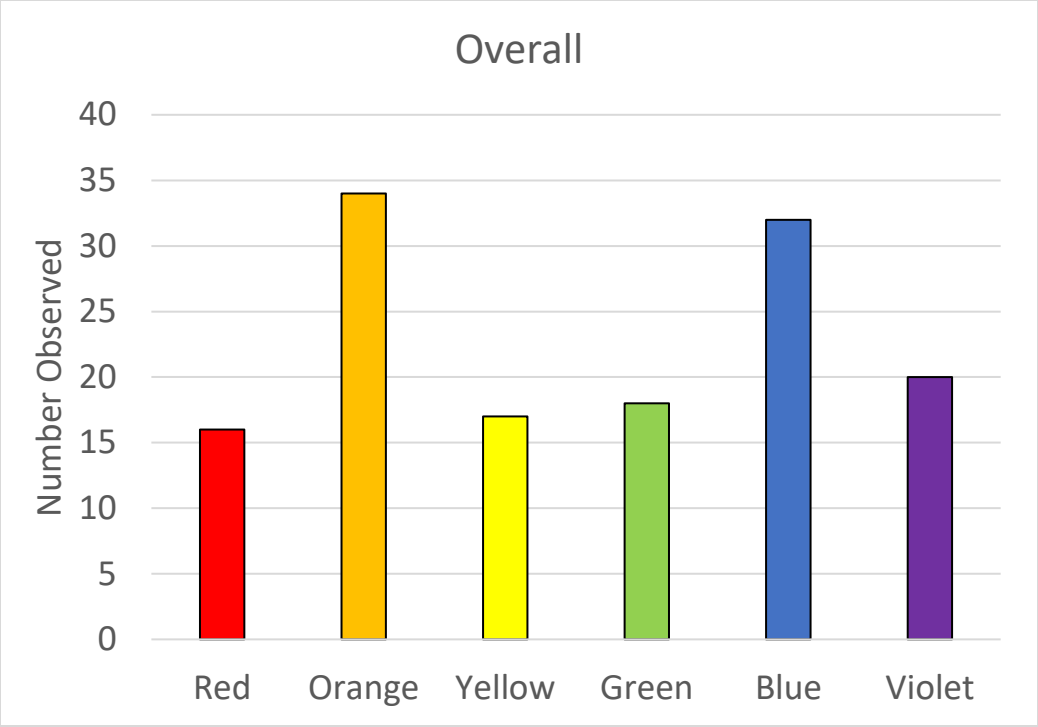

Figure 3.26. Overall frequency of FCs observed by colorblind participants. Colorblind individual saw less red and green overall when compared to trichromatic normal vision observers.

When comparing overall number of FCs observed and saturation levels of colorblind participants to normal vision participants in clear top patterns between three representative lighting conditions (red, green, and blue, Figure 3.27), blue false color observation is similar between both groups, but performance degrades for colorblind individuals under green and especially red ambient illumination.

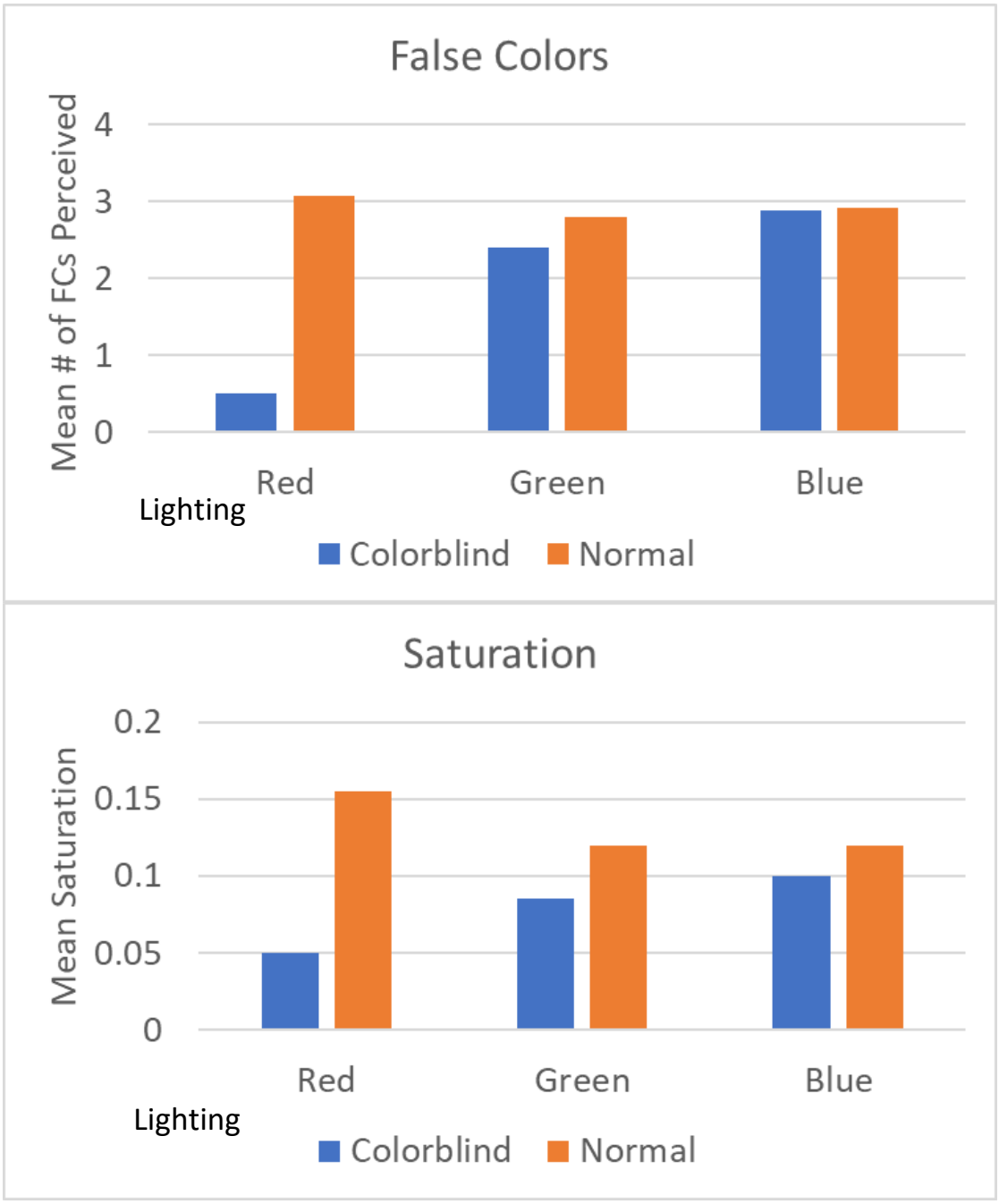

Figure 3.27. False colors (top) and saturation (bottom) perceived by colorblind compared to normal vision observers. Performance between the two samples is similar in blue lighting but degrades for colorblind individuals toward the red end of the spectrum.

Discussion

Overall, hypothesis one, that clarity of the disk pattern would enhance ability to perceive FCs was supported. Participants reported seeing more FCs as well as more saturated FCs when the top pattern was uniform rather than graded. Next, hypothesis two, that using narrowband lighting over natural broadband lighting to enhance FCs was supported. Compared to natural light, observers reported more and stronger FCs in any of the narrowband ambient lighting conditions. The specific ambient wavelength (spectral color) itself was not important as no differences were seen between the various prototypical colors. Finally, hypothesis three was only partially supported, all three top patterns were similar in number of FCs seen, but for saturation ratings, the first top pattern from the artificial spectrum top by Benham (1894) worked best in natural light and longer wavelengths, whereas the other two tops patterns showed increased saturation at shorter wavelength narrowband ambient lighting. The fourth hypothesis was supported, as colorblind individuals perceived both fewer FCs and less saturated FCs, with a marked decrease in the red wavelengths similar to that of natural light, however like the normal vision participants, colorblind individuals did see an improvement with sharper patterns and aside from red, the other narrowband wavelengths enhanced perception of false color. These findings align with previous work where colorblind individuals' results mirrored those that would be found when viewing traditional colored stimuli (Rosenblum et al., 1981; White et al., 1977). Von Campenhausen and Schramme (1995) find the same, with colorblind participants showing increased variance in matching color chips to FCs compared to normal trichromats.

Past work on ambient lighting has shown a reversal in location of false color location (Vienot & Le Rohellec, 1992), however this work used color filters that allowed a broader band of light to come through than what the current experiment's LEDs would allow. Additionally, Vienot and Le Rohellec forced participants to select chromatic responses, which fails to account for those participants who did not see false colors at all in particular locations of the Benham top.

Recent research has shown how false colors can be used to diagnose various cognitive disorders such as Alzheimer's disease (Kaubrys et al., 2016) in which those with Alzheimer's are significantly less likely to see blue false color. Such evidence further suggests that false color processing happens at the cortical and not retinal level of processing. Others have shown that retinal disease can decrease the ability to perceive false color (Pilz & Marré, 1993), but this is most likely due to reduced light to the retina, a factor already known to influence the perception of false color (Tritsch, 1992). By having a better understanding of the behavioral constraints and mechanisms surrounding the illusion of FCs, methods of testing can be improved and subsequently the usefulness of FCs in diagnosing higher order cortical visual processing dysfunction within the brain.

False colors, then, appear to be a result of processing in V1 and V2 areas with some further processing downstream in V4 (Mullen et al., 2015; Tanabe et al., 2011), with the current results pointing toward the same mechanisms. Due to the reversal of the disk direction resulting in a relocation of FCs, a timing difference as a means of explanation of false color as originally proposed by Fechner in 1838 is not tenable. Also, as the current experiment demonstrates, ambient colored illumination alone cannot

account for how many FCs are seen as well as their saturation as no significant difference was observed between the narrowband colored lighting conditions.

In summary, this experiment demonstrates that having a clear top pattern with narrowband ambient lighting produces the most optimal PIFCs in both number of FCs seen as well as saturation of observed FCs. Red is most commonly observed on the outer arc, followed by orange then green and finally blue on the innermost arc of the spinning disk. As for top pattern type, the original Benham's top pattern is optimal for most light sources and longer wavelength narrowband ambient lighting, whereas other patterns are better suited for shorter wavelength narrowband ambient lighting. These results are useful in creating demonstrations of the Benham's top to others as well as a stimulus to test for color perception in psychophysical tasks. Using these results also confirms alternative ways of creating displays of color from black and white for both normal vision and colorblind individuals.

CHAPTER 4

GENERAL DISCUSSION

Experiment 1 has found that the prototypical colors seen by both colorblind and normal vision participants are red, yellow, green and blue with the modal normal vision participants additionally identifying orange and violet as a prototypical colors. Hence, the four colors all participants perceived were red, yellow, green and blue with orange being expanded out of yellow. This is in line with the unique colored opponent process colors of the Red-Green and Yellow-Blue systems (Pridmore, 2013; Pridmore, 2020). Notably as well, these are the same four colors participants identified in experiment 3, assuming a common yellow-orange category. Experiments 2 and 3 both use these six identified prototypical colors to test for limits of temporal color processing and the potential for ambient colored illumination to enhance the false color seen in Benham's top, respectively. The experimental apparatus from the first experiment shows promise as a quick inexpensive method to efficiently determine di-, tri-, and tetrachromacy as well as provide a simple straight-forward way of testing for categories in color perception. Experiment 2 demonstrates the *where* processing system has a temporal processing threshold of around 20Hz and the *what* processing system's processing threshold is in the 2-7 Hz range. Experiment 3 found that the best conditions for seeing false color in a false color top include using a clear top pattern, narrowband ambient illumination, and the original Benham's top pattern.

Further, the four false colors commonly seen in Benham's top and supported by the analysis here indicate the universality of the universally chosen four colors in Experiment 1. These four colors again correspond to the unique opponent process colors

(Red-Green, Yellow-Blue) which is explained by a neural perspective, however the additional common categorical boundaries seen by trichromats and tetrachromats in Experiment 1 remains in question as they are not well explained by neural models or linear color theory, so seem likely to be largely a top-down phenomenon. Also of note is that the order in which false color is perceived on the top corresponds to spectral frequencies. That is, long wavelength red is most commonly observed on the outermost arc of the top, with short wavelength blue or violet on the inner arc, and the colors between being ordered similar to what would be observed in an ecologically valid stimulus of a rainbow, with the prototypical order being red-orange-green-blue. If spun in the other direction, the ordering of these colors is preserved, but in reverse locations with blue-violet being on the outermost arc of the top and red on the innermost arc. Finally, the optimal 5 Hz cycle rate of Benham's top that was used is similar to the 2-7 Hz cycle rate at which the *what* processing system tends to break down, which is also near that of the critical color fusion rate. Hence similar timing mechanisms appear to underlie both the false colors of Experiment 3 and the breakdown of color change identification of the *what* system in Experiment 2. This is again consistent with some aspect of false colors operating as a higher-level cognitive process.

The principal contribution of the present set of Experiments is to identify the prototypical color categories of bi-, tri-, and tetra-chromatic individuals, and confirm that the various timing thresholds for *What* and *Where* system processes do not appear to be dependent on wavelength. The prism color identification method introduced in Experiment 1 has the benefit that it is inexpensive, fast, and able to discriminate bi-, tri-, and tetra-chromatic individuals, and can tell us their percentages in the population. The

What versus *Where* system comparison done in Experiment 2 provides an objective method of measuring processing rates of the two neural systems that can guide potential neural models. And the False Color findings of Experiment 3 provides further insights on both the prototypical opponent-process colors and aspects of their saturation and temporal dynamics. Taken together, the three experiments provide an overview of the common categorical boundaries and temporal processing limits of human color vision.

REFERENCES

Arbab, S., Brindle, J.A., Matusiak, B.S. & Klockner, C.A. (2018). Categorisation of colour terms using new validation tools: A case study and implications. *i-Perception, 9,* 1-20.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33.

Bae, G., Olkkonen, M., Allred, S.R., & Flombaum, J.I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General,* 144, 744-763.

Backhaus, W., Kliegl, R., Werner, J.S. (1998). *Color vision: Perspectives from different disciplines.* De Gruyter.

Benham, C. E. (1894). The artificial spectrum top. *Nature*, *51*, 200-200.

Bidwell, S. (1897). On subjective colour phenomena attending sudden changes of illumination. *Proceedings of the Royal Society of London, 60*, 359-367.

Bochko, V.A. & Jameson, K.A. (2018). Investigating potential human tetrachromacy in individuals with tetrachromat genotypes using multispectral techniques. *Electronic Imaging, 14,* 1-12.

Bompas, A., Kendall, G., & Sumner, P. (2013). Spotting fruit versus picking fruit as the selective advantage of human colour vision. *i-Perception*, *4*(2), 84-94.

Boynton, R.M., Fargo, L., Olson, C.X., & Smallman, H.S. (1989). Category effects in color memory. *Color Research and Application, 14*, 229-234.

Boynton, R.M. & Olson, C.X. (1987). Locating basic colors in the OSA space. *Color Research and Application, 12,* 94-105.

Boynton, R.M., Schafer, W., & Neun, M.E. (1964). Hue-wavelength relation measured by color-naming method for three retinal locations. *Science, 146,* 666-668.

Braun, C. L., & Smirnov, S. N. (1993). Why is water blue? *Journal of Chemical Education, 70*, 612.

Campenhausen, C., Hofstetter, K., Schramme, J., & Tritsch, M. F. (1992). Color induction via non-opponent lateral interactions in the human retina. *Vision research*, *32*, 913-923.

Changzi, M.A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology Letters, 2*, 217-222.

Chen, V.J. & Cicerone, C.M. (2002). Subjective color from apparent motion. *Journal of Vision, 2*(6).

Choudhury, A. K. R. (2014). *Principles of colour and appearance measurement: Object appearance, colour perception and instrumental measurement*. Elsevier.

CIE (1926). Commission internationale de l'Eclairage proceedings, 1924. Cambridge University Press, Cambridge.

Cohen, J., & Gordon, D. A. (1949). The Prevost-Fechner-Benham subjective colors. *Psychological Bulletin, 46*, 97.

Creem, S. H., & Proffitt, D. R. (2001). Defining the cortical visual systems: "What","where", and "how". *Acta Psychologica, 107*, 43-68.

de Groot, J. J., & Van Vliet, J. A. J. M. (1986). *The high-pressure sodium lamp*. Macmillan International Higher Education.

Edwards, M., Goodhew, S.C., & Badcock, D.R. (2021). Using perceptual tasks to selectively measure magnocellular and parvocellular performance: Rationale and a user's guide. *Psychonomic Bulletin & Review,* 1-22.

Eisen-Enosh, A., Farah, N., Burgansky-Eliash, Z., Polat, U. & Mandel, Y. (2017). Evaluation of critical flicker-fusion frequency measurement methods for the investigation of visual temporal resolution. *Scientific Reports, 7*, (15621).

Fechner, G. T. (1838) Über eine Scheibe zur Erzeugung subjektiver Farben. *Annalen der Physik und Chemie 45*, 227 -232

Fink, M., Ulbrich, P., Churan, J., & Wittmann, M. (2006). Stimulus-dependent processing of temporal order. *Behavioral Processes, 71*, 344-352.

Finlayson, G., Mackiewicz, M., Hurlbert, A., Pearce, B. & Chrichton, S. (2014). On calculating metamer sets for spectrally tunable LED illuminators. *Journal of the Optical Society of America, 31*(7), 1577-1587.

Finnegan, J.M. & Moore, B. (1894). The artificial spectrum top. *Nature, 51,* 292-293.

Fortner, B. & Meyer, T.E. (1998). *Number by colors: A guide to using color to understand technical data.* Springer-Verlag New York, Inc.

Gegenfurtner, K.R. & Riger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology, 10,* 805-808.

Goldstein, E.B. (2007). *Sensation and perception* (7th ed.). Wadsworth: Thompson.

Gordon, N. (1998). Colour Blindness. *Public Health, 112,* 81-84.

Green, D.G. (1969). Sinusoidal flicker characteristics of the color-sensitive mechanisms of the eye. *Vision Research, 9*(5), 591-601.

Goodale, M.A., Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience. 15*(1), 20–5.

Goodale, M.A. & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience, 15*(1), 20-25.

Guild, J. 1932. The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London,* Series A. 230, 149-187.

Habili, N. & Oorloff, J., Scyllarus (2015). *From Research to Commercial Software,* In Proceedings of the ASWEC 2015 24th Australasian Software Engineering Conference, Adelaide, Australia, pp 119-122, September 2015.

Harris, A. H., Fernandes-Taylor, S., & Giori, N. (2012). "Not statistically different" does not necessarily mean "the same": The important but underappreciated distinction between difference and equivalence studies. *JBJS, 94,* e29.

Holloway, S.R., McBeath, M.K. & Macknik, S.L. (2012). Color helps isolate dorsal stream contribution to shape-recognition task. *Journal of Vision, 12*(9), 1311.

Holloway, S.R. & McBeath, M.K. (2013). Independent objective timing tests designed to measure processing rates of dorsal and ventral visual systems, *Journal of Clinical and Experimental Neuroscience, 1*(2),15-20.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory, 7*(6), 418.

Jarvis, J.R. (1977). On Fechner-Benham subjective colour. *Vision Research, 17,* 445-451.

Jameson, K.A. (2009). Human potential for tetrachromacy. *Glimpse: The Art+ Science of Seeing, 2,*82-91.

Jameson, K.A. & Highnote, S.M., & Wasserman, L.M. (2001). Richer color experience in observers with multiple photopigment opsin genes. *Psychonomic Bulletin &*

Review, 8,
244-261.

Jameson, K.A. & Komarova, N.L. (2009). Evolutionary models of color categorization. I. Population categorization systems based on normal and dichromat observers. *Journal of the Optical Society of America A, 26,* 1414-1423.

Jameson, K.A., Satalich, T.A., Joe, K.C., Bochko, V.A., Atilano, S.R., & Kenney, M.C. (2020). *Human color vision and tetrachromacy.* Cambridge University Press.

Jones, L. A. (1943). Historical background and evolution of the colorimetry report. *Journal of the Optical Society of America. 33*(10): 534–43.

Jordan, G., Deeb, S.S., Bosten, J.M., & Mollon, J.D. (2010). The dimensionality of color vision in carriers of anomalous trichromacy. *Journal of Vision, 10,* 1-19.

Jordan, G. & Mollon, J. (1993). A study of women heterozygous for colour deficiencies. *Vision Research, 33,* 1495-1508.

Jordan, G. & Mollon, J. (2019). Tetrachromacy: The mysterious case of extra-ordinary color vision. *Current Opinion in Behavioral Sciences, 30,* 130-134.

Kaubrys, G., Bukina, V., Bingelytė, I., & Taluntis, V. (2016). Perception of Fechner illusory colors in Alzheimer disease patients. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 22,* 4670.

Kellogg, J. M., & O'Shea, G. (2007). *Ocular dominance and subjective color perception: a study using the Fechner-Benham Visual Illusion* (Doctoral dissertation, Colorado State University. Libraries).

Kenyon, G. T., Hill, D., Theiler, J., George, J. S., & Marshak, D. W. (2004). A theory of the Benham Top based on center–surround interactions in the parvocellular pathway. *Neural networks, 17,* 773-786.

Kessels, R. P., Postma, A., & de Haan, E. H. (1999). P and M channel-specific interference in the what and where pathway. *Neuroreport, 10,* 3765-3767.

Kong, X., Murdoch, M. J., Vogels, I., Sekulovski, D., & Heynderickx, I. (2019). Perceived speed of changing color in chroma and hue directions in CIELAB. *JOSA A, 36,* 1022-1032.

Kraus, N. & Nicol, T. (2005). Brainstem origins for cortical ‘what’ and ‘where’ pathways in the auditory system. *Trends in Neuroscience, 28,* 176-181.

Krynen, R.C. (2021). Rainbow stripes: Categorical perception of color as a tool for testing tetrachromacy**.** Manuscript in progress.

Lad, S.S., Hurley, R.A., & Taber, K.H. (2020). Temporal processing: Neural correlates and clinical relevance. *The Journal of Neuropsychiatry and Clinical Neurosciences,* 32, A6-108.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*, 259-269.

Landis, C. (1954). Determinants of the critical flicker-fusion threshold. *Physiological Reviews*, *34*, 259-286.

Le Rohellec, J., & Vienot, F. (2001). Interaction of luminance and spectral adaptation upon Benham subjective colours. *Color Research & Application*, *26*(S1), S174-S179.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(2), 467-477.

Lucas, P.W., Dominy, N.J., Riba-Hernandez, P., Stoner, K.E., Yamashita, N., Loría-Calderón, E., Petersen-Pereira, W., Rojas-Durán, Y., Salas-Pena, R., Solis-Madrigal, S., Osorio, D., & Darvell, B.W. (2003). Evolution and function of routine trichromatic vision in primates. *Evolution, 57,* 2636-2643.

Maloney, L.T. & Wandell, B.A. (1986). Color constancy: A method for recovering surface spectral reflectance. *Journal of the Optical Society of America, 3*(1), 29-33.

Marois, R. & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Science, 9*(6), 296-305.

Masri, R.A., Grünert, U., & Martin, P.R. (2020). Analysis of parvocellular and magnocellular visual pathways in human retina. *Journal of Neuroscience, 40,* 8132-8148.

McBeath, M. K. (1990). *Velocity of apparent motion in three-dimensional* space (Doctoral dissertation). Stanford University, Stanford, CA.

McBeath, M., & Krynen, R. C. (2015). Velocity of the Human Stadium or Mexican La Ola Wave: Systematic Variations Due to Type and Direction. *Journal of vision*, *15*, 748-748.

McBeath, M. K., Tang, T. Y., & Shaffer, D. (2018). The geometry of consciousness. *Cognition and Consciousness, 64*, 207-215.

McKeeff, T. J., Remus, D. A., & Tong, F. (2007). Temporal limitations in object processing across the human ventral visual pathway. *Journal of neurophysiology, 98*, 382-393.

Merigan, W. H., Katz, L. M., & Maunsell, J. H. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *Journal of Neuroscience*, *11*, 994-1001.

Milne, B. F., Toker, Y., Rubio, A., & Nielsen, S. B. (2015). Unraveling the intrinsic color of chlorophyll. *Angewandte Chemie International Edition*, *54*, 2170-2173.

Milner, A.D. & Goodale, M.A. (2008). Two visual systems re-viewed. *Neuropsychologica*, *46*(3).

Mishkin, M. & Ungerleider, L.G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioral Brain Research, 6*, 57–77.

Mullen, K. T., Chang, D. H., & Hess, R. F. (2015). The selectivity of responses to red-green colour and achromatic contrast in the human visual cortex: an fMRI adaptation study. *European Journal of Neuroscience*, *42*, 2923-2933.

Nakano, Y. & Kaiser, P.K. (1992). Color fusion and flicker fusion frequencies using tritanopic pairs. *Vision Research*, *32*(8), 1417-1423.

Natural Resources Conservation Service (2021, May 9). *Exploring soil colors.* https://www.nrcs.usda.gov/wps/portal/nrcs/detail/wi/soils/?cid=NRCSEPRD1370419

Neitz, M., Kraft, T.W., & Neitz, J. (1998). Expression of L cone pigment gene subtypes in females. *Vision Research, 38*, 3221-3225.

Nishiyama, Y. (2012). Benham's top. *Osaka Keidai Ronshu*, *62*, 87-94.

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences*, *25*(1), 73-144.

Picton, T. (2013). Hearing in Time: Evoked potential studies of temporal processing. *Ear and Hearing, 34*, 385–401.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in cognitive sciences, 1*(2), 56-61.

Rauschecker, J. P. (2018). Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex, 98*, 262-268.

Rosenblum, K., Anderson, M. L., & Purple, R. L. (1981). Normal and color defective perception of Fechner-Benham colors: Implications for color vision theory. *Vision research, 21*, 1483-1490.

Schramme, J. (1992). Changes in pattern induced flicker colors are mediated by the blue/yellow opponent process. *Vision Research. 32*(11): 2129–2134.

Shapiro, K. (2001). The limits of attention: Temporal constraints in human information processing. New York, NY, US: Oxford University Press.

Shepard, R.N. & Cooper, L.A. (1992). Representation of colors in the blind, color-blind, and normally sighted. *Psychological Science, 3,* 97-104.

Sill, C.W. (1961). Transmittance spectra of color filters. *Analytical Chemistry,* 33, 1584-1587.

Simonson, E., & Brozek, J. (1952). Flicker fusion frequency: Background and applications. *Physiological reviews, 32*, 349-378.

Skelton, A.E., Catchpole, G., Abbott, J.T., Bosten, J.M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences of the United States of America, 114,* 5545-5550.

Tanabe, H. C., Sakai, T., Morito, Y., Kochiyama, T., & Sadato, N. (2011). Neural correlates and effective connectivity of subjective colors during the Benham's top illusion: a functional MRI study. *Cerebral Cortex, 21*, 124-133.

Thones, S., von Castell, C., Iflinger, J. & Oberfeld, D. (2018). Color and time perception: Evidence for temporal overestimation of blue stimuli. *Scientific Reports, 8*(1688).

Tritsch, M. F. (1992). Fourier analysis of the stimuli for pattern-induced flicker colors. *Vision research, 32*, 1461-1470.

Truss, C. V. (1955). Chromatic flicker fusion frequency as a function of chromaticity difference. *Journal of the Optical Society of America, 47*(12), 1130-1134.

Uchikawa, K. & Shinoda, H. (1996). Influence of basic color categories on color memory discrimination. *Color Research and Application, 21,* 430-439.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*, 1181-1186.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*, 598-607.

Vienot, F. & Le Rohellec, J. (1992). Reversal in the sequence of the Benham Colors with a change in the wavelength of illumination. *Vision Research, 32*(12), 2369-2374.

von Campenhausen, C. (1973). Detection of short time delays between photic stimuli by means of pattern induced flicker colors (PIFCs). *Vision research*, *13*(12), 2261-2272.

von Campenhausen C., Hofstetter K, Schramme J., Tritsch M. F., (1992) Color induction via nonopponent lateral interactions in the human retina. *Vision Research, 32*, 913 -923

von Campenhausen, C. & Schramme, J. (1995). 100 years of Benham's top in colour science. *Perception*, *24*, 695-717.

Wagstaff, K., Cardie, C., Rogers, S., & Schroedl, S. (2001, June). Constrained k-means clustering with background knowledge. In Icml (Vol. 1, pp. 577-584).

White, C. W., Lockhead, G. R., & Evans, N. J. (1977). Multidimensional scaling of subjective colors by color-blind observers. *Perception & Psychophysics*, *21*(6), 522-526.

Willson, M.F. & Whelan, C.J. (1990). The evolution of fruit color in fleshy-fruited plants. *The American Naturalist, 136,* 790-809.

Witzel, C. & Gegenfurtner, K.R. (2018). Color perception: Objects, constancy, and categories. *Annual Review of Vision Science, 4*, 475-499.

Wright, W. D. 1929. A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*. 30, 141-164.

Wright, W. D. 1930. A re-determination of the mixture curves of the spectrum. *Transactions of the Optical Society*. 31, 201-218.

APPENDIX A

ISHIHARA TEST PLATES USED IN EXPERIMENT 1

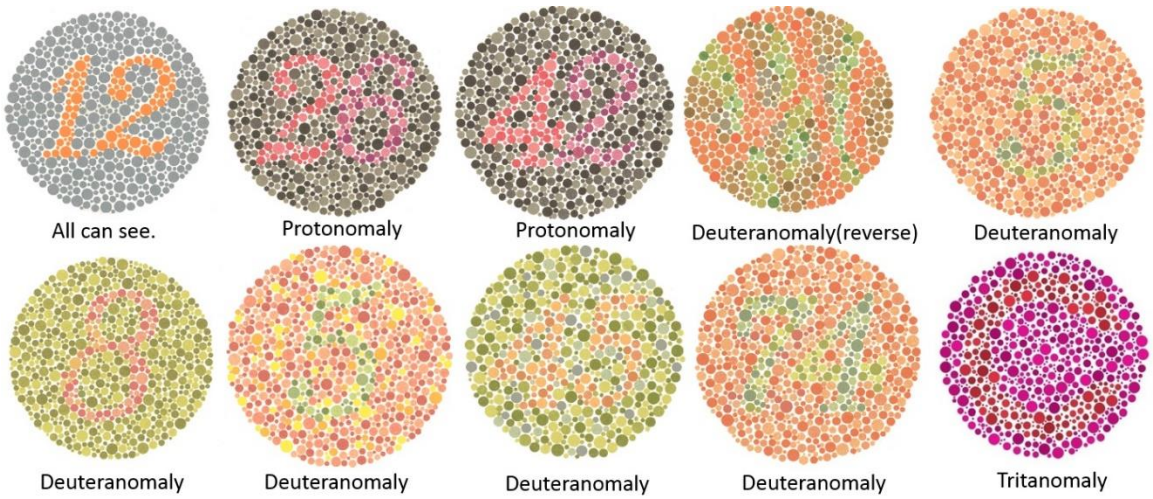

Figure A-1. Examples of Ishihara plates used to test for colorblindness.

APPENDIX B

COLOR-AID CORP STIMULI USED IN EXPERIMENT 3

Figure B-1. Photograph of the 60 possible color choices participants could select. Hue, saturation and lightness are represented in these color choices.

Table B-1.
Table of Color-Aid Corp Values and the Labeling System Used For The Current Experiment.

Red			Orange			Yellow		
R1	R2	R3	O1	O2	O3	Y1	Y2	Y3
R-P1-1	Rc-hue	Rc-T2	O-P1-1	Ro-Hue	RO-T2	Y-P1-1	YG-Hue	YG-T1
R4	R5	R6	O4	O5	O6	Y4	Y5	Y6
R-P2-2	R-hue	R-T2	O-P2-2	O-Hue	O-T2	Y-P2-2	Y-Hue	Y-T2
R7	R8	R9	O7	O8	O9	Y7	Y8	Y9
R-P2-1	Rw-hue	Rw-T2	O-P2-1	YO-Hue	YO-T2	Y-P2-1	Yc-Hue	Yc-T2

Green			Blue			Purple		
G1	G2	G3	B1	B2	B3	V1	V2	V3
G-P1-1	Gw-hue	Gw-T2	B-P1-1	Bc-Hue	Bc-T2	V-P1-1	Bv-hue	Bv-T2
G4	G5	G6	B4	B5	B6	V4	V5	V6
G-P2-2	G-Hue	G-T2	B-P2-2	B-Hue	B-T2	V-P2-2	V-hue	V-T2
G7	G8	G9	B7	B8	B9	V7	V8	V9
G-P2-1	Gc-hue	Gc-T2	B-P2-1	Bw-hue	Bw-T2	V-P2-1	Rv-hue	Rv-T2

1	2	3	4	5	6
WHITE	Gray 2.5	Gray 4.5	Gray 6.5	Gray 8.5	BLACK

Table B-2.
Experiment Labels and Their Corresponding Values in CIE 1931 Color Space.

	x	y
R1	0.433	0.2907
R2	0.5181	0.2668
R3	0.4507	0.2795
R4	0.3713	0.3099
R5	0.5544	0.2886
R6	0.5067	0.2887
R7	0.3909	0.3077
R8	0.583	0.309
R9	0.5251	0.3064
O1	0.4683	0.3611
O2	0.5847	0.3249
O3	0.513	0.3314
O4	0.4319	0.3557
O5	0.5979	0.3331

O6	0.4878	0.3521
O7	0.4568	0.3601
O8	0.5481	0.3739
O9	0.4666	0.3748
Y1	0.4079	0.4121
Y2	0.2926	0.5276
Y3	0.3048	0.465
Y4	0.3912	0.3943
Y5	0.472	0.4445
Y6	0.3994	0.4136
Y7	0.4162	0.4045
Y8	0.4115	0.4803
Y9	0.3846	0.4218
G1	0.263	0.4271
G2	0.2081	0.4873
G3	0.2509	0.4492
G4	0.2873	0.3737
G5	0.2003	0.4425
G6	0.2234	0.4205
G7	0.2781	0.387
G8	0.1958	0.3804
G9	0.2083	0.384
B1	0.2206	0.2295
B2	0.187	0.2399
B3	0.2033	0.2682
B4	0.2652	0.2892
B5	0.1903	0.2109
B6	0.2065	0.2479
B7	0.2492	0.2715
B8	0.1977	0.1861
B9	0.2103	0.2148
V1	0.2717	0.2452
V2	0.2431	0.2018
V3	0.2393	0.2048
V4	0.2898	0.2759
V5	0.2952	0.2104
V6	0.2788	0.2201
V7	0.2886	0.2744
V8	0.3476	0.2217
V9	0.3236	0.2463

Black	0.32	0.3255
Gray 2.5	0.3199	0.3259
Gray 4.5	0.317	0.3242
Gray 6.5	0.3186	0.3251
Gray 8.5	0.3203	0.327
WHITE	0.3187	0.329

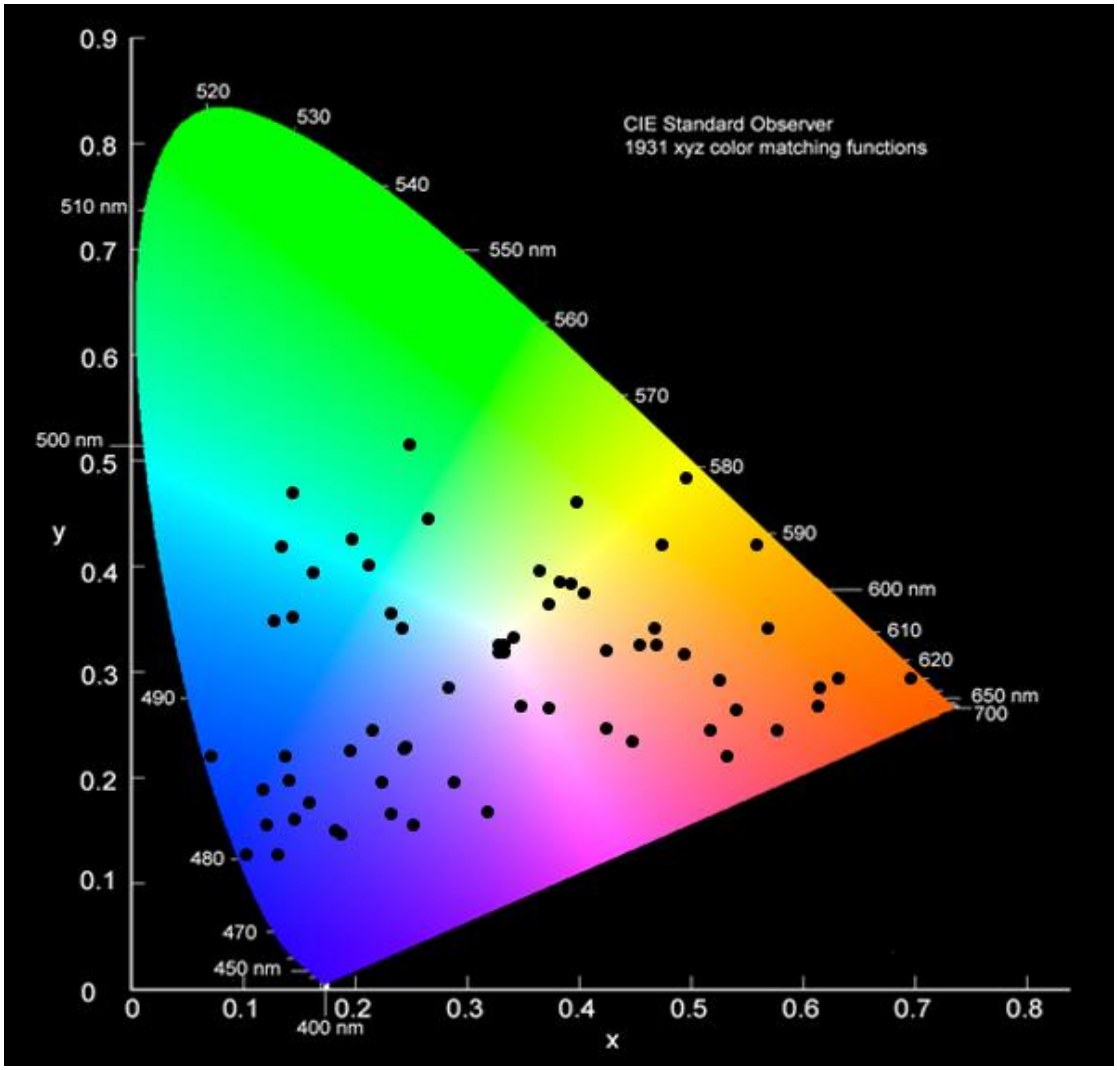

Figure B-2. Plotted points of all 60 possible color choices in CIE 1931 space.

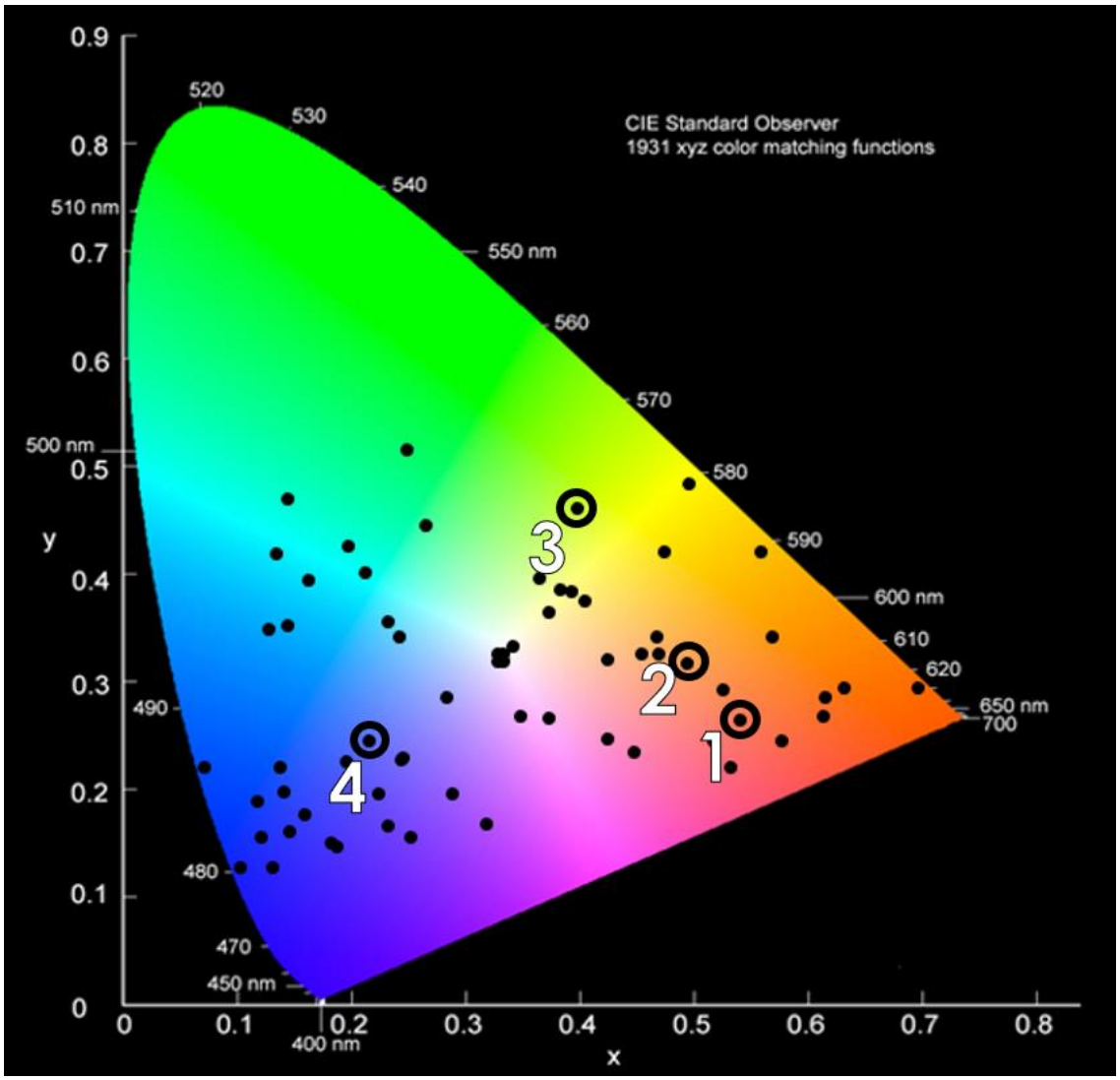

Figure B-3. Plotted prototypical colors seen by normal vision participants in numbered order of arc (1 is outermost arc, 4 is innermost for reference).

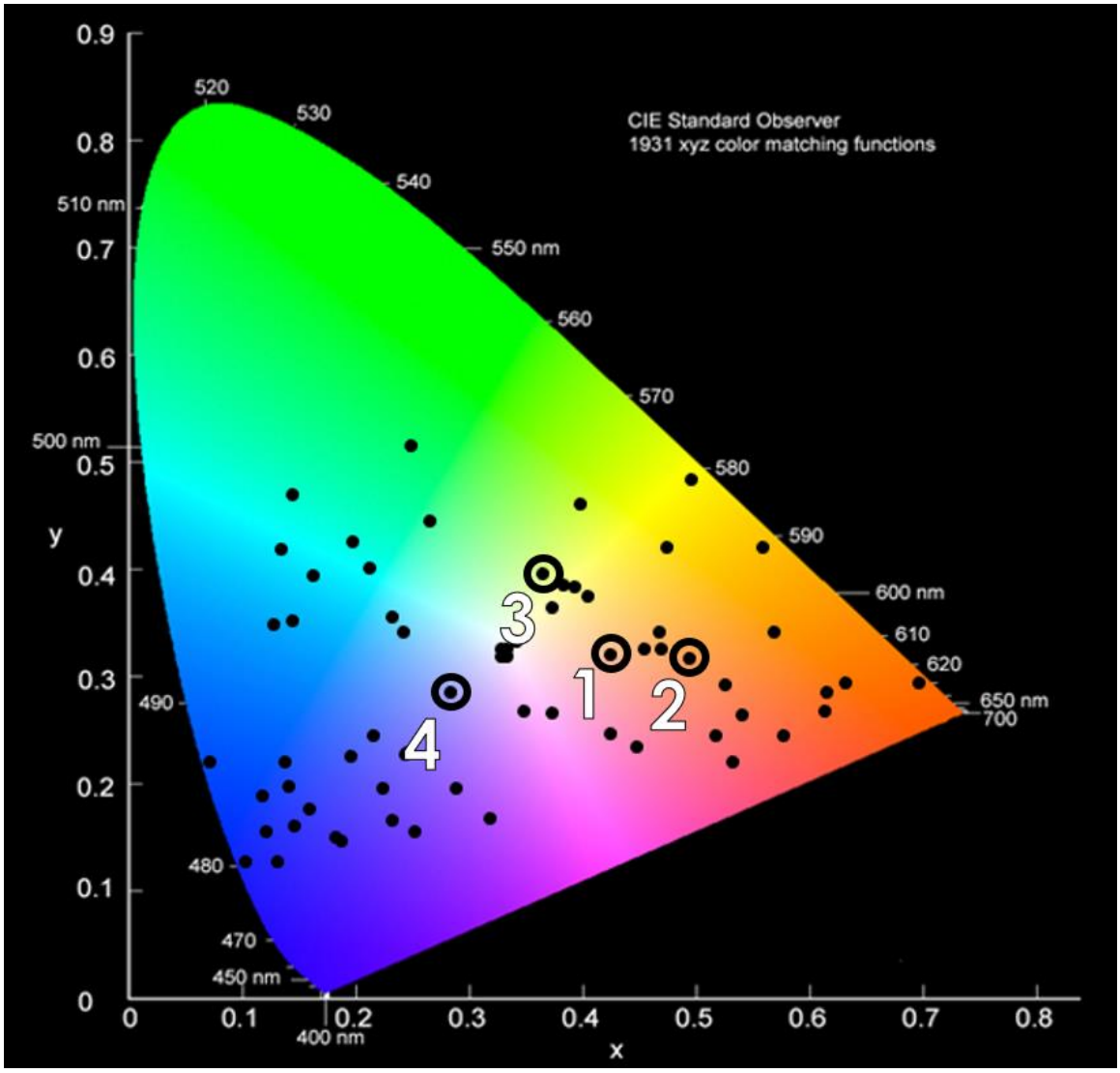

Figure B-4. Plotted prototypical colors seen by colorblind participants in numbered order of arc (1 is outermost arc, 4 is innermost for reference).

APPENDIX C

IRB APPROVAL DOCUMENT

ASU Knowledge Enterprise Development

EXEMPTION GRANTED

Michael McBeath
CLAS-NS: Psychology
480/965-8930
Michael.McBeath@asu.edu

Dear Michael McBeath:

On 6/4/2019 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Temporal Color Perception and Seeing False Colors
Investigator:	Michael McBeath
IRB ID:	STUDY00010242
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• HRP-503a Temporal Color Study.docx, Category: IRB Protocol; • Temporal Color Perception Recruitment Flyer.pdf, Category: Recruitment Materials; • Temporal Color Consent Form Draft 6.pdf, Category: Consent Form; • Stimuli and Tools.pdf, Category: Technical materials/diagrams; • Demographic Information TCS.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 6/4/2019.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc: Richard Krynen
Jacob Price
Richard Krynen