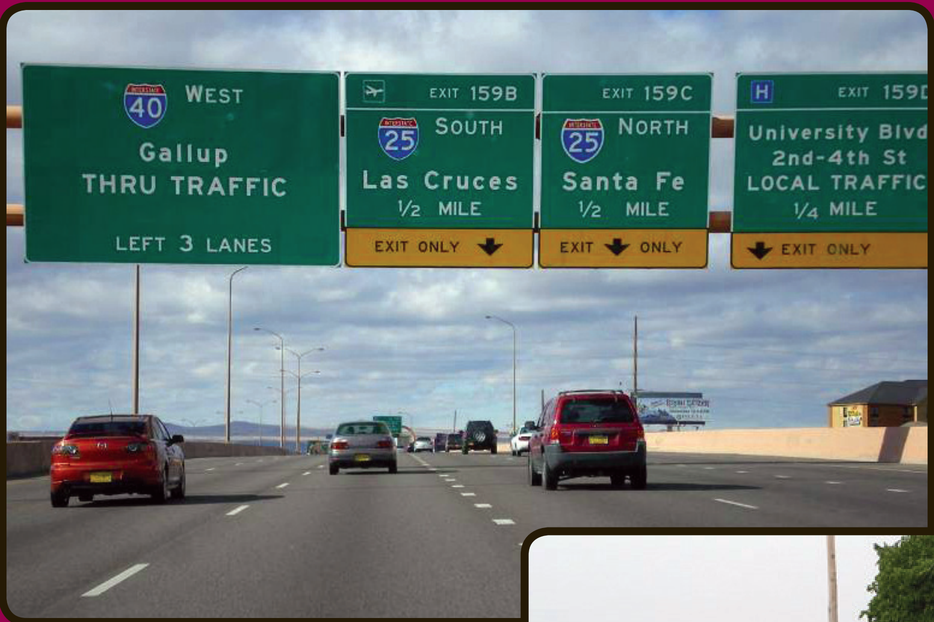
Daytime Color Appearance of Retroreflective Traffic Control Sign Materials

PUBLICATION NO. FHWA-HRT-13-018 APRIL 2013



U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

Retroreflective materials used for street and highway traffic control signs were developed to increase nighttime visibility by reflecting a maximum amount of light back from the headlights of a vehicle to the eyes of a driver. The retroreflective properties of these sign materials increase the variability of photometric measurements taken from the materials both in the laboratory and in the field. In addition, the retroreflective properties of these sign materials may affect their color appearance when viewed by drivers under daylight conditions.

This report describes a research study conducted to determine physical measurements of the chromaticity and luminance of retroreflective sign materials by means of instruments and to determine perceptual measurements of the color appearance (hue, apparent saturation, and brightness) of these materials as judged by a group of human observers. Comparisons are presented between physical measurements made in the laboratory and in the field and between these physical measurements and the psychophysical determination of color appearance obtained from a sample of 17 observers. These comparisons have implications for the specification of allowed color ranges for retroreflective sign materials.

This report will be of interest to Federal, State, and local agencies concerned with specifying and maintaining the color properties of retroreflective traffic control signs, to sign material manufacturers, and to researchers studying the visibility of signs as related to highway safety.

Monique R. Evans
Director, Office of Safety
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-13-018	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Daytime Color Appearance of Retroreflective Traffic Control Sign Materials		5. Report Date April 2013
		6. Performing Organization Code:
7. Author(s) John A. Molino, Jason F. Kennedy, Pascal A. Beuse, C. Cameron Miller, Wendy Davis, and Carl K. Andersen		8. Performing Organization Report No.
9. Performing Organization Name and Address Science Applications International Corporation (SAIC) 8301 Greensboro Drive M/S T1-12-3 McLean, VA 22102		10. Work Unit No.
		11. Contract or Grant No. DTFH61-08-C-00006
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Safety Research and Development Turner-Fairbank Highway Research Center 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Covered Final Report, January 2007–May 2010
		14. Sponsoring Agency Code HRDS-05
15. Supplementary Notes The FHWA Contracting Officer's Technical Representative (COTR) was Christopher Monk, HRDS-30.		

16. Abstract

Photometric measurements of the daytime chromaticity and luminance of retroreflective sign materials were made both in the laboratory and in the field. These instrument measurements were compared with daytime perceptual judgments of color properties made by human observers in the field. Hue, saturation, and brightness were determined for four different types of retroreflective sheeting materials and one diffuse material at four quadrants of each color area for six different colors specified in the Code of Federal Regulations for use on street and highway traffic control signs.

Overall, participants' responses aligned well with the instrument measurements. The introduction of retroreflective properties reduced both the apparent saturation and the brightness of all the colors. The results for the white, green, and blue color areas showed distinct color separations for retroreflective sign materials. The results for the red, orange, and yellow areas showed less color separation but little overlap among contiguous colors in the perceptual color space.

This outcome has implications for determining the shape and size of the color areas used to specify the colors that may be employed on traffic control signs. While the present color areas seem to adequately support the daytime perception of the color of retroreflective sign materials for the six colors tested, special attention needs to be paid to the red/orange and orange/yellow color boundaries in any future revisions to the Federal Highway Administration (FHWA) color areas. The results of the present experiment indicate that there is no pressing need to modify the FHWA color areas, but improvements might be made in future iterations so as to enhance separation of those color areas that are more difficult to discriminate.

17. Key Words Daytime color appearance, Retroreflective signs, Hue scaling, Saturation scaling, Brightness scaling, Color boxes, Chromaticity, Luminance		18. Distribution Statement No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.	
19. Security Classif. (of this report): Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages: 71	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

PROBLEM STATEMENT

Daylight measurements of the color of retroreflective materials used for traffic control signs are generally more variable than daylight color measurements of uniform diffuse surfaces.[1] Retroreflective materials have spherical or prismatic elements that direct light back in non-uniform ways. These materials are primarily designed to reflect light from the headlights of a vehicle back toward the vehicle to improve nighttime visibility for the driver. The color of the material under daylight viewing conditions is usually of secondary concern.[1] The reproducibility of daylight color measurements of retroreflective materials in the field and in the laboratory shows considerable variability, with different measuring instruments yielding different results.[2] Further, instrument measurements of chromaticity often do not correspond to perceived color judgments made by human observers.[2] Understanding and reducing these inconsistencies are important to the Federal Highway Administration (FHWA) in defining the size and shape of the color areas used to specify colors for traffic control signs. These color areas, or color boxes, are incorporated into the Code of Federal Regulations (CFR) as color specifications for retroreflective sign and pavement marking materials.[3]

BACKGROUND

The color of an object depends on, among other parameters, the spectral properties of the illuminant. International Commission on Illumination (CIE) standard illuminant D65, which is representative of average daylight in the northern hemisphere, has a correlated color temperature (CCT) of 6,500 K. Illuminant D65 produces a substantially different spectrum than illumination from CIE standard illuminant A, which has a CCT of 2,856 K and is intended to represent typical tungsten-filament lighting. Illuminant A is traditionally used to make nighttime measurements of retroreflective materials as illuminated by the tungsten-halogen headlights of a vehicle. Therefore, in the United States, Federal regulations give color specifications of retroreflective sign materials that are different for nighttime and daytime conditions.[3]

For retroreflective materials, FHWA daylight chromaticity coordinates defining the acceptable color regions for the six colors used in the present experiment are plotted in the CIE 1931 *x, y* chromaticity diagram in figure 1.[4] The corresponding CIE chromaticity coordinates for the same six colors are shown in figure 2.[4] Important differences exist between the two sets of daytime color coordinates for retroreflective sheeting. The CIE coordinates specify a smaller color area for white and a somewhat smaller color area for orange and display a wider hue angle separation (more blank space between lines of constant hue) between the red, orange, and yellow areas. Although the relationship between color coordinates and hue discrimination by human observers is not precise, this wider separation is an indication of less potential color confusion in the CIE system. One of the purposes of the present experiment was to examine whether the smaller separation in the FHWA system is adequate to support color discrimination of retroreflective sign materials by human observers.

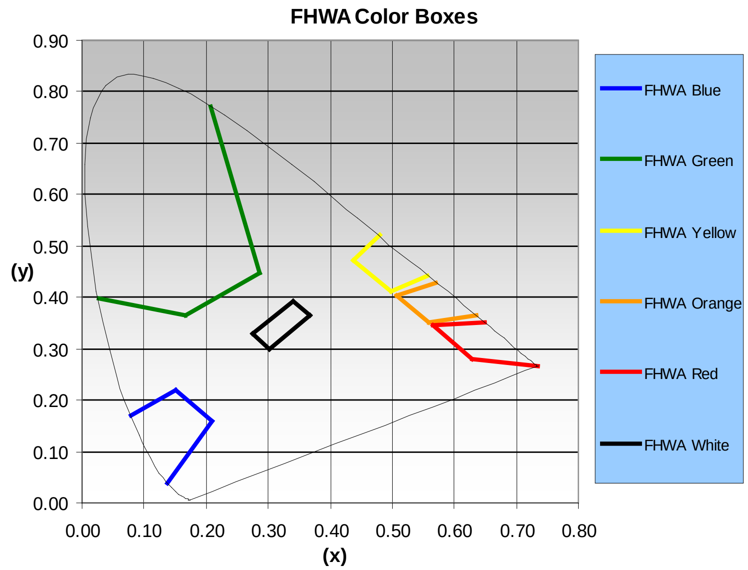


Figure 1. Graph. FHWA daylight chromaticity coordinates for various colors of retroreflective materials plotted in the CIE 1931 color space.

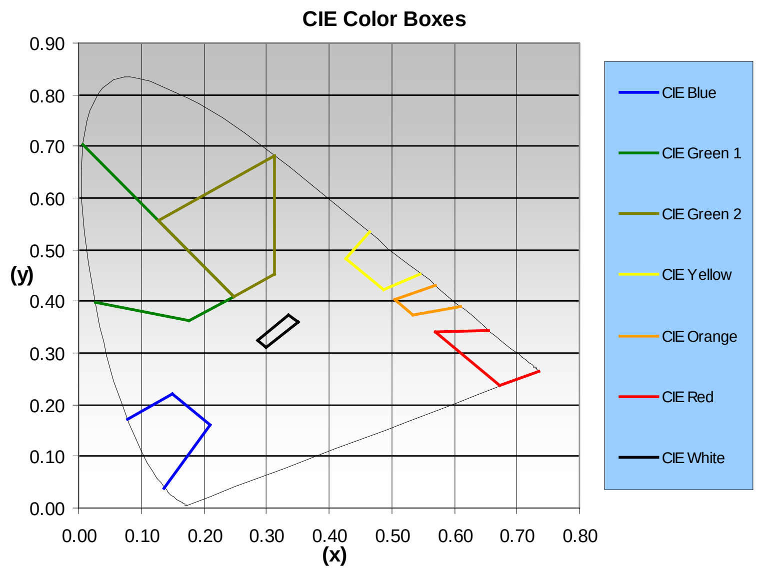


Figure 2. Graph. CIE daylight chromaticity coordinates for various colors of retroreflective materials plotted in the CIE 1931 color space.

The retroreflective properties of sign materials significantly improve the visibility of traffic control signs during nighttime viewing, but they also affect their color appearance and brightness during daylight viewing.(1) The present study uses four different retroreflective materials and a standard of diffuse reflection to investigate the effects of various retroreflective properties on the chromaticity and luminance of reflected light under natural daylight and simulated daylight (illuminant D65) conditions. The study compares these instrument measurements with the daytime color appearance of the materials as judged by a group of human observers. Although it is difficult to relate luminance to brightness in a precise manner, the study evaluates the degree of correlation existing between measured luminance and apparent brightness to gain insight into the relative brightness of these various retroreflective materials as they might appear to drivers under daylight conditions.

Abramov et al. developed a technique for specifying color appearance by human observers that has considerable advantages over traditional means such as tristimulus colorimetry.[(5)] The method does not require precise equipment or controlled viewing conditions and thus lends itself readily to field applications. The technique is based upon direct perceptual hue and apparent saturation scaling. Research participants give percentages of their sensations based on four unique hue names: red, yellow, green, and blue. They also give a separate achromatic percentage for apparent saturation. The results of these direct scaling determinations of color appearance are expressed on a uniform appearance diagram (UAD), an opponent color diagram composed of two orthogonal axes (red-green and yellow-blue). Abramov et al. related the results of their measurements to those of other color specification systems. In one example, the researchers were able to derive traditional measures of discriminability from UADs. In another example, wavelength discrimination could be predicted from the distances among stimuli on a UAD. In fact, the authors claim that the UAD method can yield discriminability data that are at least as metrically uniform as the 1976 CIE diagram.[(5)]

Gordon et al. expanded on this research and presented further evidence of the robustness of the technique.[(6)] The researchers provided detailed reasons for the choice of the four color names and claimed that participants do not need special training to use these terms, since the unique hue components refer to internal standards. They suggested using an arcsine transformation to reduce non-uniformities in variance due to having bounded scales between 0 and 100 percent. Gordon et al. explored the effects of long-term stability over several months, context provided by preceding stimuli, stimulus range, experience giving hue scaling judgments, and language. They found that the hue and apparent saturation scaling method was robust in the face of these effects and produced accurate and reliable data. They noted that about 5 percent of participants did not use the percentage scales appropriately and their data had to be excluded. The method described by Gordon et al. formed the basis for the color appearance ratings collected in the present study.[(6)]

Jacobs and Johnson conducted a study of the color appearance of different retroreflective pavement marking products.[(7)] Although the study was not directly concerned with retroreflective sign sheeting materials, the results may be relevant since pavement marking materials also involve both color and retroreflectivity. In their experiment, seven color-normal human observers made color rating judgments on a scale of 1 to 5 from white to yellow. The observers sat in a stationary motor vehicle and rated pavement marking samples at viewing distances from 39 to 118 ft (12 to 36 m). While the researchers did not use the method of hue and apparent saturation scaling, they did show a correlation of their rating scale with chromaticity measurements of different pavement marking products. The authors found significant differences in the color ratings for the various pavement marking materials. Daytime and nighttime color determinations were not the same. Some yellow pavement marking materials were rated as yellow under daylight illumination but white when illuminated with the low beam headlights of the test vehicle at night.[(7)]

Thomas-Meyers et al. presented findings indicating that a specific range of chromaticities acceptable for pavement marking colors could be useful for both daytime and nighttime viewing conditions.[(8)] The authors measured hue and apparent saturation ratings from 34 research participants who viewed colored pavement marking stripes, either yellow (center line) or white (edge line), on a background of pavement color presented on a computer display. Although there were minor differences, the color regions that were reliably judged as yellow or white were similar under both daytime and nighttime contrast conditions. Older and younger participants gave similar color appearance judgments, but color-deficient participants judged the colors somewhat differently, especially at night. Chromaticity coordinates for proposed Ohio standards for yellow and white pavement markings were consistent with color appearance data for the white color, but the yellow color needed to be shifted toward higher saturation values. When in-use pavement markings were compared to the proposed standards, the white markings were consistent, but some of the aged yellow pavement markings appeared to be white. Although Thomas-Meyers et al. employed a relatively large sample of 34 participants, they used less realistic viewing conditions. The stimuli were presented on a computer display, and the study concentrated on the color appearance of pavement markings rather than highway traffic signs.[(8)]

In 2007, Davis and Miller conducted a pilot field study for FHWA on the daytime color appearance of retroreflective materials used for highway traffic signs. The pilot field study led to the research described in the present report.[(9)] The researchers made physical measurements of the chromaticity coordinates and luminance for a small sample of FHWA color specifications (selected from those shown in figure 1) both in the laboratory and in the field. Five naïve observers used the hue and apparent saturation rating method described by Gordon et al. to rate the appearance of the colors.[(6)] Six different types of retroreflective materials were employed, along with a standard diffuse white reflector. The hue and apparent saturation results showed less saturation for both the yellow and orange retroreflective samples when compared with the diffuse reflector with the same colors. Based on limited data, the authors tentatively concluded that the instrument measurements were generally consistent with the color appearance judgments of human observers in terms of both hue and saturation.

RESEARCH APPROACH

The present study investigated the color appearance of retroreflective materials used for traffic control signs. The experiment compared field judgments of perceived hue, apparent saturation, and brightness made by human observers with instrument measurements of chromaticity, saturation (derived from the CIE 1976 $L^*a^*b^*$ (CIELAB) calculations), and luminance made with spectroradiometers in the laboratory and in the field.[(10)] While the term *chroma* is used to describe the concept of saturation in CIELAB, the term *saturation* is used interchangeably in this report. Perceptual and physical measurements were made for 120 color samples.

Although brightness is difficult to measure and to correlate with physical measurements made by instruments, brightness was included in the investigation because it could yield further insight into the perception of drivers regarding retroreflective sign material. This study expanded on the methodology of Gordon et al. by adding a brightness scale to the procedure.[(6)] The present study also included a larger participant sample (n = 17) than was used in earlier studies.[(5,6)] In addition, the study employed samples of real highway traffic sign materials viewed under actual daylight field conditions, instead of small spots of light presented in a laboratory as in the originating studies. Gordon et al. reported somewhat low within-participant variability; however, participants in the earlier studies were experienced, having participated in several experiments, and rather homogenous, coming from an academic environment.[(6)] It was uncertain how age and gender could affect variability for a sample of naïve participants recruited from the general driving public. The present study included a more heterogeneous group of participants, including both male and female as well as a wide age range.

CHAPTER 2. METHODOLOGY

This chapter describes the overall design of the experiment, the participants, the materials, and the procedures used in the study. It covers both instrument measurements and perceptual judgments.

STIMULI

The stimuli for the present study consisted of 120 test samples. Of these samples, 24 consisted of a standard of diffuse reflectance with a reflective index of 98 percent paired with one of a variety of color and neutral density filters. Measurements of these samples established a reference condition for color measurement. The other 96 samples consisted of 4 different types of white retroreflective sign material covered with color and neutral density filters selected to match the chromaticity and luminance factor values obtained for the diffuse samples. The four different sheeting materials tested were ASTM types III, VIII, IX, and proposed type XI.[11] Of the 13 colors specified for use in traffic control signs, the 6 colors tested were: red, green, blue, yellow, orange, and white. There were four variants for each of the six colors, with each variant approximating one of the four corners of the color area that defines that color in CFR Title 23.[3]

INSTRUMENT MEASUREMENTS

The present experiment investigated daylight chromaticity and luminance properties of sample sign materials by means of measurements taken with spectroradiometric instruments both in the laboratory and in the field. The experiment compared these instrument measurements with the perceptual judgments of color and brightness given by human observers in the field.

Measurements were taken using two different spectroradiometers and a spectrocolorimeter. Laboratory measurements were made with a PR-715 SpectraScan® spectroradiometer and a Hunter LabScan® XE spectrocolorimeter. Field measurements were made with a PR-650 SpectraScan® spectroradiometer.

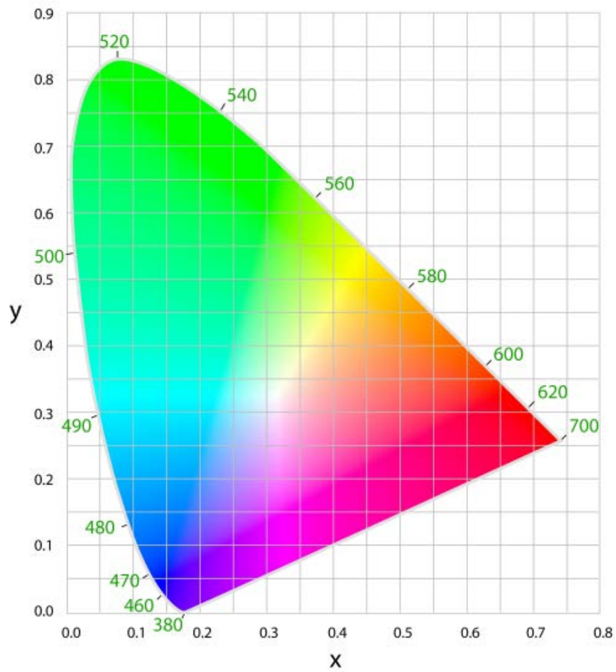
The purpose of using two different laboratory measurement procedures was to compare the values obtained using a spectroradiometer and a bench-top spectrocolorimeter. The values obtained by the PR-715 spectroradiometer were then compared to the field measurements made using the PR-650 spectroradiometer. This chain of instrument measurements was intended to provide a basis for correlating the observational conditions to standard laboratory conditions under which measurements are made for quality assurance. It was also important to gain insight into differences in laboratory measurements used to determine compliance with standards and perceptual judgments of color, apparent saturation, and brightness.

Laboratory measurements of the 120 color samples and the standard diffuse white reflector were taken with the PR-715 at the National Institute of Standards and Technology on an optical bench using procedures specified by ASTM E1349.[12] Measurements were made for each sample using a 0.5-degree-high by 1.5-degree-wide rectangular aperture. A 0-degree/45-degree geometry was employed, where the illuminating source was at 0 degrees (normal incidence) and the measurement instrument was at 45 degrees, relative to the surface of the sample being tested. The PR-715 was placed approximately 7 ft (2 m) from the sample such that an area of 2.92 by 0.689 inches (74.1 by 17.5 mm) was measured. The illuminating source was a 1,000-W FEL CC8 tungsten-halogen lamp operated at approximately 3,200 K and placed 7 ft (2 m) from the sample. The illuminant was corrected by calculation to match illuminant D65. A single measurement of the diffuse reflector and each color sample was made. The spectroradiometric values obtained were used to calculate tristimulus values (X, Y, Z) for each sample, which were converted to chromaticity coordinates (x, y) and a luminance factor (Y). The samples were also measured in the laboratory at the FHWA research site with a Hunter LabScan® XE. The LabScan® XE uses annular illumination at 45 degrees with measurements at 0 degrees. Eight measurements were made with the LabScan® XE and averaged. The sample was removed and replaced on the measurement port before each measurement. This procedure was intended to account for the smaller area sampled by the LabScan® XE compared to the PR-715.

Field physical measurements were initially made at the FHWA research site during the color appearance trials with human observers in fall of 2007. These measurements were taken at a distance of 32.5 ft (9.9 m) on the same horizontal plane as the stimuli. The instrument's line of sight was approximately 5 degrees from the observers' line of sight. This technique was used to approximate the participants' viewing angle (0 degrees) without interfering with their observations. The measurements were made using a 1-degree spot aperture under diffuse illumination produced by the daytime sky. Because of the fast pace of stimulus presentation, the researchers were not able to measure every sample presentation. Photometric and colorimetric measurement was taken every 10th trial of the color appearance judgments. Since the stimuli were presented to the participants in different random orders for each session, this sampling procedure resulted in an uneven number of measurements for the 120 color samples. Chromaticity (x, y), luminance (Y), and CCT determinations were made for each sample. In addition to the physical measurements made in conjunction with the color appearance judgments, physical measurements of the standard diffuse white reflector with no color filters were made before and after each block of 60 color appearance trials. These additional measurements were used to track lighting changes due to varying sky conditions (passing clouds, high overcast, etc.) within experimental sessions and across sessions conducted on different days.

Since the field luminance measurements taken with the PR-650 were sampled only every 10th perceptual trial, the sample sizes were small and unequal. To obtain a more statistically reliable sample, the field chromaticity and luminance measurements were reproduced outdoors 1 year later, in fall of 2008, with the same instrument (PR-650). The 120 color samples were measured seven times at a 0-degree angle of regard. This 0-degree angle for the corroborate measurements represented true normal orientation but was not substantially different from the approximately 5-degree angle used in fall of 2007.

The CIE 1931 x, y chromaticity coordinates for both the laboratory and field physical measurements were converted into the CIELAB color space for comparison with UADs from the human observers (see figure 3 and figure 4).[4] The CIELAB color space is based on an opponent color concept, which is similar to the opponent color processes of human vision.[10] The CIELAB color space coordinates were rotated 90 degrees clockwise to more closely correspond with the layout of the perceptual UADs. Such a rotation resulted in a reversal of the positive and negative vertical axis. Since the emphasis in the present experiment was on perceptual measurements, the customary UAD layout was employed as the reference.



Numbers around periphery represent wavelengths in nm.

Figure 3. Graph. CIE 1931 *x*, *y* chromaticity diagram.

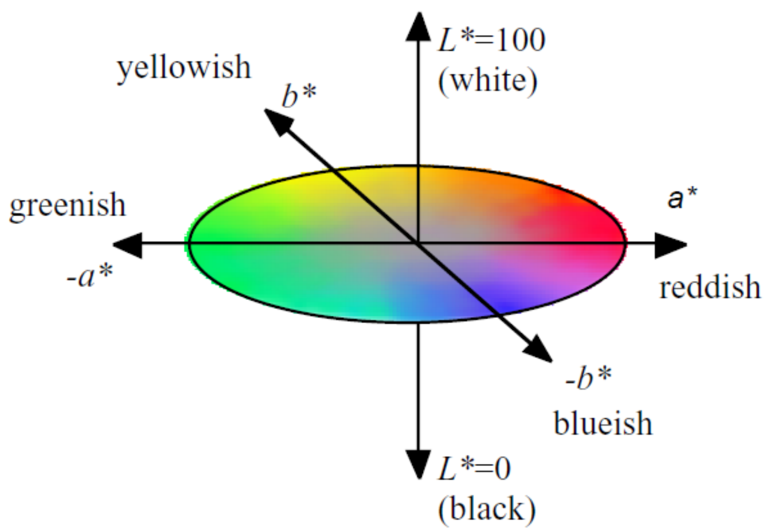


Figure 4. Illustration. CIELAB color space.

PERCEPTUAL MEASUREMENTS

The perceptual experiment examined the responses of participants to colored retroreflective sign materials under daytime lighting conditions using two methodologies. A rating scale technique was used to determine the hue, apparent saturation, and brightness responses for individual color samples. A ranking technique was used to determine relative brightness responses for a subset of color samples.

Participants

Seventeen people were recruited from the Washington, DC, metropolitan area. All participants were adults (at least 18 years of age), were licensed drivers, had visual acuity (corrected or uncorrected) of at least 20/40 in their best eye, and had normal color vision. Table 1 shows the median and mean age by gender and age category. There were nine men (six younger, three older) and eight women (five younger, three older). Participants were assigned to one of three groups, which produced two groups of six and one group of five. Each group was assigned a day of the week (Tuesday, Wednesday, or Thursday) for coming to the research facility each week for the duration of the experiment. The experiment and its procedures were approved by an institutional review board. The participants received an honorarium for their services.

Table 1. Participant characteristics by age and gender category.

Category	Number	Mean (Years)	Median (Years)	Age Range
Younger (18–64 years)	11	22	23	19–25
Older (65+ years)	6	71	68	66–83
Men	9	40	25	20–83
Women	8	38	24	19–68

Environment and Materials

The research was conducted outdoors on the grounds of the FHWA Turner-Fairbank Highway Research Center (TFHRC) in McLean, VA. The experiment was conducted during September, October, and early November 2007 in daylight, with the middle of each 4-h session occurring when the sun was at the meridian. The sessions were conducted only on clear, partly cloudy, or high overcast days in relatively bright sunlight conditions. Testing was conducted in an area that was away from distractions such as vehicular and pedestrian traffic. Figure 5 depicts the testing environment and relative location of the equipment used in the experiment.

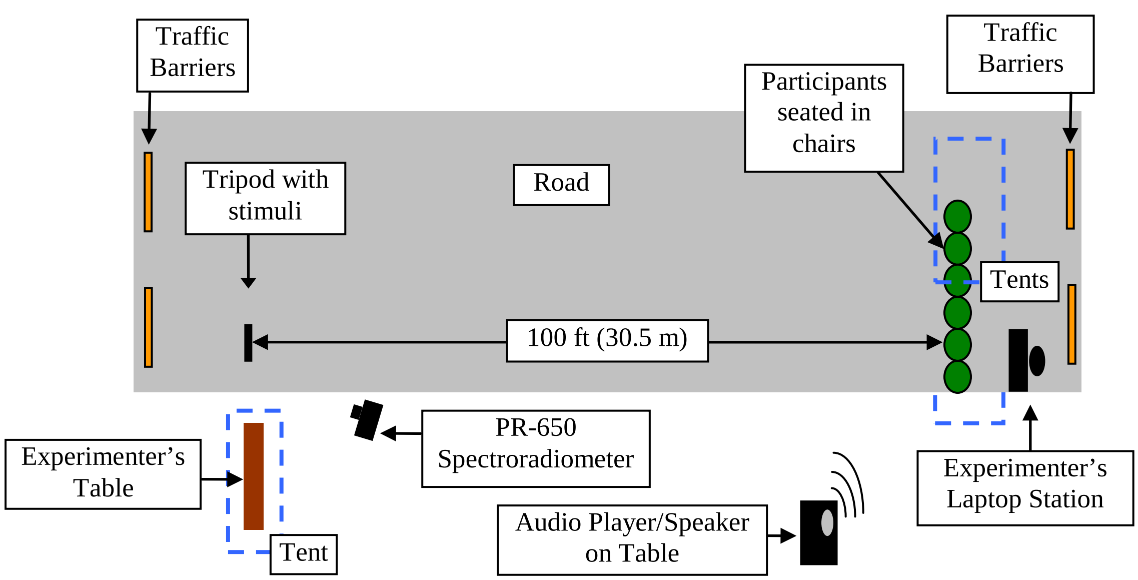


Figure 5. Illustration. Plan view of the experimental setup (not to scale).

Participants observed the stimuli and then marked their judgments on preprinted response sheets mounted on clipboards. The timing of judgment trials was accomplished by prerecorded verbal commands. An audio file was played aloud through the speakers to cue the participants when to respond to the individual color samples during the rating task. The audio cue was used to keep an observation time pace of 10 s per stimulus.

Additional equipment included tripods and a laptop computer. Tripods were used to hold the color samples for the individual rating task, the subset of color samples for the ranking task, and the portable PR-650. The laptop computer was used to record data from the PR-650.

The generally accepted daylight legibility distance for a standard 30-by-30-inch (76-by-76-cm) STOP sign with 10-inch (25-cm)-high lettering is 400 ft (122 m).[13] The visual stimulus samples of color and sheeting combinations were 7.5 inches (19.1 cm) square, quarter scale relative to a standard STOP sign. Therefore, participants viewed the samples from a quarter-scaled distance of 100 ft (30.5 m). This setup is not to imply that color recognition of traffic signs is limited to the legibility distance but, rather, reflects the distance at which a driver may acquire information regarding a particular sign. The color and shape of a traffic sign should be unambiguous at the legibility distance to ensure that all informational aspects of a sign reinforce each other.

The color samples were mounted at a height of 5 ft (1.5 m), the minimum mounting height of a rural STOP sign.[13] The samples were tilted slightly forward to minimize specular reflections, as is typically done in highway traffic sign installations. Figure 6 shows a standard red STOP sign with mounting characteristics similar to those replicated in the experiment. Note the green foliage background exhibited in this view. The colored samples in the experiment were viewed on a background of natural grass, bushes, and trees with a fence in the far distance.

Figure 6. Photo. Standard retroreflective STOP sign in a typical application.

Procedure

The experiment was conducted outdoors at the edge of a roadway on the TFHRC grounds. The roadway was blocked by traffic barriers during testing. Participants sat in chairs in a single row under two shade tents and observed the color samples from a different seating position for each half-day block of trials. The seating positions were assigned randomly at the beginning of each testing session. Individual color samples were mounted on a tripod located at the side of the roadway, 100 ft (30.5 m) from the row of seated participants. A third shade tent, located to the side of the tripod containing the sample, was used to cover the table containing the racks of color samples. One experimenter stood near this third tent to change the color samples on the tripod between trials. The second tripod holding the PR-650 was located slightly to the left of the participants' line of sight, 32.5 ft (9.91 m) from the color sample tripod. The audio loudspeaker system was located on a small table between the color sample tripod and the row of participant chairs. A second experimenter sat in a chair behind the participants to monitor the laptop data collection from the PR-650. Figure 7 shows the outdoor experiment setup.

Figure 7. Photo. Outdoor experiment setup on the grounds of TFHRC.

Each group participated in four 4-h sessions over a period of four or five weeks. The first session was somewhat longer to allow time for paperwork and training. Examples of the training materials used to acquaint the participants with the required responses are in appendix A.

Day 1 of the Experiment

Before data collection began, participants read and signed an informed consent form and were administered a visual acuity and color vision test. Participants' visual acuity was assessed with a standard wall-mounted Snellen chart. Color vision was assessed with an Ishihara color deficiency test book. Participants were required to have at least 20/40 vision, uncorrected or corrected, in at least one eye and to have normal color vision.

Following the vision testing and initial administrative paperwork, participants were taken to the outdoor testing facility to begin training. The primary training took place at the beginning of the first day for each participant group and was scheduled to take 30–45 min. Participants were given detailed instructions regarding the four tasks that they were to perform during the study: hue scaling, apparent saturation scaling, brightness scaling, and brightness ranking. The formal verbal instructions were read aloud, and additional training consisted of an explanation of the concepts of hue, apparent saturation, and brightness along with diagrams and charts. A practice exercise was administered, and ample time was allotted for questions throughout the training. The formal verbal instructions as well as the supplemental training materials are in appendix A.

The participants judged each color sample by means of a modified method of hue and apparent saturation scaling.[6] In this method, each participant rated the color of the sample stimulus in terms of the percentages of four unique hues (red, yellow, green, and blue), such that the total percentage added to 100 percent. Participants could assign 100 percent of the hue to one color, or they could use pairs of the basic hues. Participants were restricted from pairing red with green and yellow with blue; otherwise, all other pairs were allowed. Each participant also rated the apparent saturation (colorfulness) of each sample stimulus on a separate 0–100 percent scale. See appendix A for instructions given on the hue and apparent saturation scaling method.

An additional brightness scale, expressed as a percentage ranging from 0 to 100 percent, was added to the aforementioned method. Verbal instructions were given for this perceptual brightness scale (see appendix A). Thus, each presentation of a color stimulus sample received three or four perceptual rating scores: one or two for the percentage of the four basic hues, one for apparent saturation (colorfulness), and one for brightness. These measurements constituted the fundamental color appearance judgments given by the participants to the 120 color stimulus samples. Each day, participants viewed the set of 120 color samples twice, once in the morning and once in the afternoon, in a different random order. Participants viewed the color samples in blocks of 60 stimuli, with a 5-min break between each block.

The participants were also asked to perform a separate perceptual brightness ranking task. Two colors, red and yellow, were used for this task. Five samples with the same color chromaticity point (red or yellow) were presented simultaneously on a tripod. The five samples represented the same color filter with the four different retroreflective sheetings and the diffuse white reflector material. The participants ranked the five samples from dimmest to brightest. These additional perceptual brightness ranking judgments were made following each block of 60 individual trials for the hue, apparent saturation, and brightness scaling tasks. Participants saw the red and yellow sample sets separately twice each day (in different left-to-right positions) for a total of eight times during the course of the study. Table 2 shows the typical daily schedule for participants in the study.

Table 2. Typical daily experimental schedule.

Task	Duration (min)
Hue, Saturation, and Brightness Rating—Block 1	30
Short Break	5
Brightness Ranking—Set 1	5
Short Break	5
Hue, Saturation, and Brightness Rating—Block 2	30
Short Break	5
Brightness Ranking—Set 2	5
Lunch Break	30
Hue, Saturation, and Brightness Rating—Block 3	30
Short Break	5
Brightness Ranking—Set 3	5
Short Break	5
Hue, Saturation, and Brightness Rating—Block 4	30
Short Break	5
Brightness Ranking—Set 4	5
Total Time	200

Days 2–4 of the Experiment

Participants returned to TFHRC three more times over the course of several weeks. Before testing began each day, the experimenters briefly reviewed the scaling and ranking procedures. Following the review, participants repeated the four data collection blocks (two morning, two afternoon) that were conducted on day 1, except the stimuli were presented in a different random order for each block. This procedure continued for days 2–4. After data collection was completed on day 4, participants were debriefed, offered an opportunity to ask questions, and paid for their participation.

The primary task in the perceptual portion of the experiment involved 8 repetitions of hue, apparent saturation, and brightness scaling judgments for 120 color samples, a total of 960 scaling judgments made by each participant. Since there were 17 participants, the entire experiment yielded 16,320 scaling judgments for the primary task. The secondary task in the experiment yielded 272 brightness rankings of one color area coordinate for the yellow and red color samples (136 rankings per color). Each ranking involved 5 brightness levels, making for 1,360 total ranking scores. These brightness ranking determinations supplemented the brightness rating measurements by employing a different method to determine similar color properties.

CHAPTER 3. RESULTS

This chapter presents the results of the physical instrument measurements using the laboratory and field spectroradiometers. Human response results are also presented, including data from the hue, apparent saturation, and brightness rating tasks as well as the brightness ranking task.

CHROMATICITY AND HUE/APPARENT SATURATION

Instrument Measurement Results

The measurements taken with the spectroradiometers were used to calculate tristimulus values (X, Y, Z), which were then transformed into the $L^*a^*b^*$ color space. The resulting values were plotted on CIELAB plots with axes that were rotated 90 degrees clockwise. CIELAB plots have two axes ($-a/+a$, $-b/+b$). The $-a/+a$ axis represents the approximate opponent colors of green and red. The $-b/+b$ axis represents the approximate opponent colors of blue and yellow. While the opponent color axes in these plots are not precisely related to the named colors, there is a strong approximate relationship. The $+a$ axis may not represent a pure red, but it is indicative of a color that is close to red. Most of the physical color measurements made in the experiment, whether in the laboratory or in the field, are expressed in terms of these CIELAB plots that have been rotated to more closely align with the UADs used to represent the perceptual color and apparent saturation judgments of the participants.

Figure 8 shows the mean laboratory and field color measurements of the white diffuse reflector plus color filters. These measurements were intended to serve as a reference against which to assess the effects of the field measurement geometry as compared to standard laboratory measurements. However, the white diffuse reflector with color filters displayed a much greater reduction in saturation than the various retroreflective samples when comparing field measurements to laboratory measurements. While the reason for this is unknown, it may have been due to the manner in which the stimuli were constructed. Plexiglas® sheets with color and neutral density filters were placed in front of the white diffuse reflector to create the diffuse color samples, while color and neutral density filters were adhered directly onto samples of the white retroreflective sheeting. Daylight entering at the edges of the Plexiglas® sheet and interreflections between the back surface of the Plexiglas® and the front surface of the white diffuse reflector may have been responsible for the measured reduction in color saturation. In any case, the reduction in saturation resulted in an inability to use the diffuse measurements as a link between laboratory and field measurements.

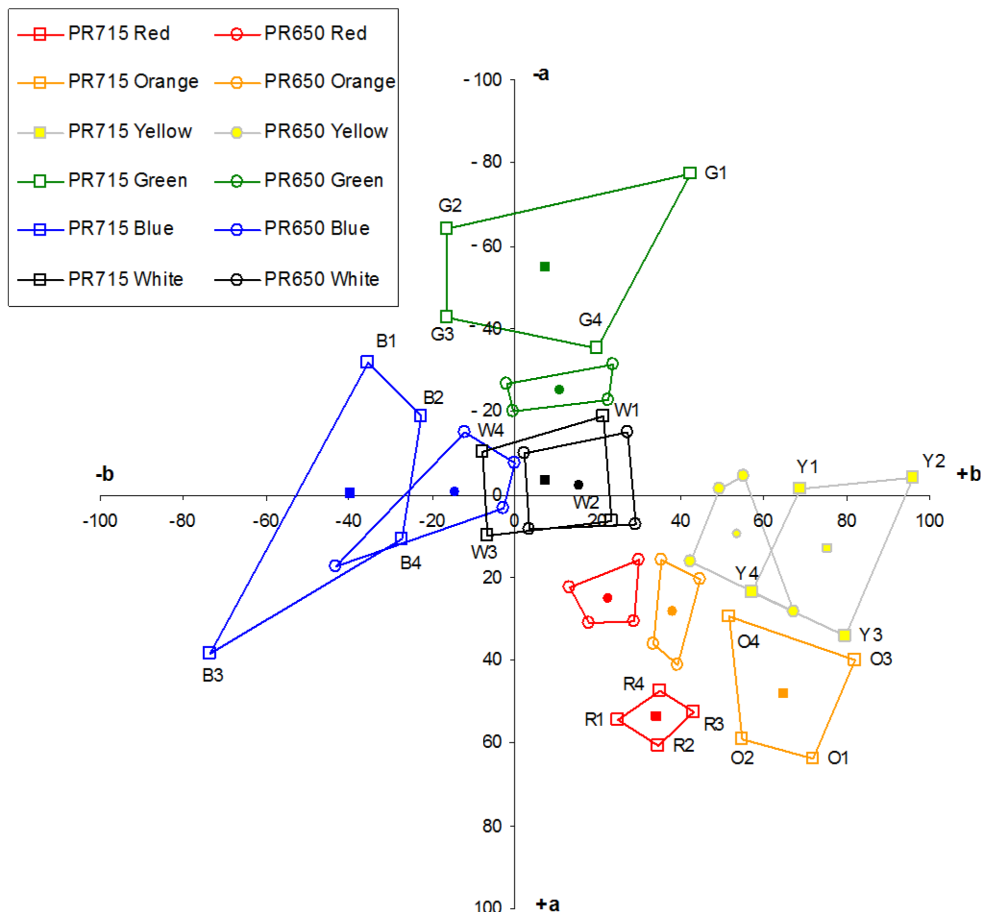


Figure 8. Graph. Mean laboratory (PR-715) and field (PR-650) physical color measurements of the white diffuse reflector with color filters.

The labels in figure 8 identify the corners of the laboratory color areas. Each label indicates the color (red [R], yellow [Y], orange [O], green [G], blue [B], and white [W]) and filter number (1, 2, 3, and 4). Due to space considerations, the labels are only provided in this graph. However, the corners of the color areas stay in the same relational orientation for all the physical measurements. Therefore, the labels in figure 8 can be referred to when viewing subsequent CIELAB plots. Since the original field measurements were made only on every 10th trial during the determination of color ratings, not all samples were measured and some samples were measured several times. Thus, the more comprehensive field chromaticity and luminance measurements made 1 year after the perceptual experiment are presented in figure 8.

Since most of the color areas represent four-sided figures, the terms *color boxes* and *color areas* are used interchangeably. In the laboratory data from the PR-715 shown in figure 8, the color areas are typically well shaped and oriented. The exception is the red box, which is the smallest of all the boxes and is not aligned on the +*a* axis but skewed toward the +*b* axis. This illustrates the difficulty in comparing color measurements with subjective color assessments. The CIELAB +*a* axis does not align with the red axis of a UAD. The other CIELAB axes are oriented closer to the named color axes (yellow, green, and blue) of a UAD. The centroid for the blue box is almost exactly on the -*b* axis. The centroid for the green box is slightly shifted toward the +*b* axis, and the centroid for the yellow box is shifted somewhat toward the +*a* axis. The centroid for the white box is also shifted toward the +*b* axis. Since the *a*,*b* color axes cannot be specifically related to named colors, these shifts may not directly relate to color changes observed by human observers and should be regarded primarily in terms of orientation in the CIELAB color space. Thus, the only major unexpected outcome for the PR-715 measurements is the relatively small size of the red color box.

The field measurements made with the PR-650 show all of the previously discussed characteristics, with two significant exceptions. First, all of the colors measured in the field, except for white, are less saturated than their laboratory counterparts. Second, the red color box had a slight shift in hue toward the +*b* axis, so that it slightly overlapped the orange color box in terms of hue angle. The hue angle is the angle of a constant hue line that radiates out from the origin of the CIELAB plot from white toward any saturated color. A given hue angle represents all colors of the same hue but different saturation. The shift in the red color box is an unexpected and unexplained result.

Figure 9 shows the mean laboratory and field physical color measurements averaged over the four retroreflective sheeting types. The standard errors of the mean for the field measurements in both figure 8 and figure 9 ranged from 1.6 to 2.3; so mean differences greater than 5 scale units are likely to be statistically significant. When compared with figure 8, figure 9 reveals much less reduction in saturation.

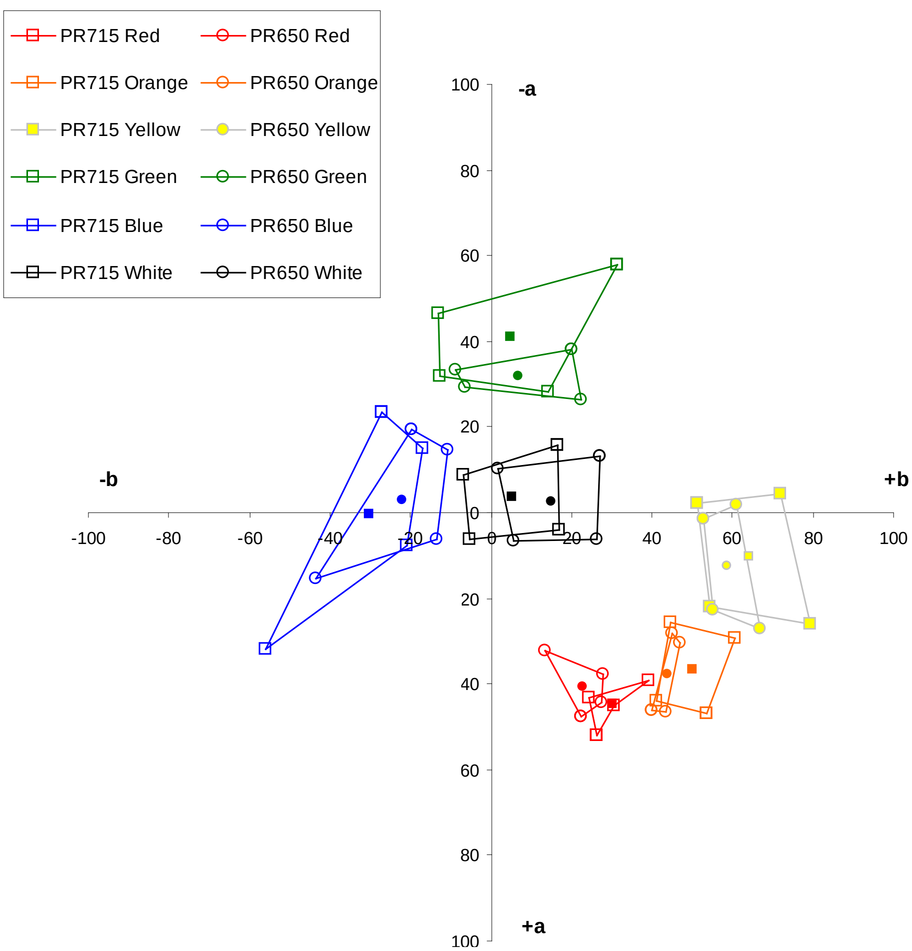


Figure 9. Graph. Mean laboratory (PR-715) and field (PR-650) physical color measurements averaged over four retroreflective sheeting types.

The PR-715 measurements of the retroreflective materials are closer to the origin (the point of zero saturation) than the measurements of the diffuse samples, and the distance between the contiguous red and orange boxes is shorter. The resulting smaller separation between the red and orange color boxes may make them more difficult to discriminate. In fact, a hypothetic constant hue line between the O2 and R3 data points indicates a slight overlap in the measured red and orange color boxes. The R4 corner of the red box has a greater level of uncertainty due to stray light concerns at long wavelengths for the PR-715.

The complete set of field measurements taken 1 year after the experiment was closely correlated to the limited set of measurements taken during the perceptual assessments (see figure 30 in appendix B). So, the results analyzed and presented are based on the more complete field measurement data set taken in 2008. The PR-650 color boxes show the same trend as the laboratory measurements, but with significantly less reduction in saturation (see figure 9). The color boxes measured in the field are in relatively the same shape and orientation as the corresponding color boxes measured in the laboratory. For field measurements of retroreflective sheeting materials, the outer boundary of each color box, except white, was shifted toward the center of the diagram relative to the laboratory measurements. The white color box was shifted slightly along the +*b* axis. The orange box moved closer to the red box and is considerably compressed along the *b* axis. The centroid of the red color box fell on the same hue line as was measured in the laboratory, indicating consistency of the spectroradiometer measurements. The absence of a shift in the hue line for the field measurements of the red retroreflective materials results in the red color remaining somewhat separated from the orange color box.

The PR-715 laboratory measurements of the O2 and R3 data points indicate a slight overlap of the red and orange color boxes, while the PR-650 field measurements indicate a slight separation between the color boxes. Taking into account the uncertainty of the measurements, this result indicates the potential for some degree of perceptual confusion at the extreme edges between the orange and red colors used on retroreflective signs. The probability of accurate discrimination between the yellow and the orange color boxes appears to be relatively better than between orange and red. The hue lines defining the long wavelength side of the yellow box and the short wavelength side of the orange box do not overlap. There is a larger separation in terms of hue angle between the blue and green color boxes and between the green and yellow color boxes, indicating that those colors of retroreflective materials are likely to produce less color confusion for people with normal color vision.

Figure 10 shows the laboratory color measurements taken with the LabScan® XE and averaged over the four sheeting types. The results are similar to the laboratory measurements taken with the PR-715. The major differences are in the red, orange, and yellow boxes. Specifically, these colors demonstrate a greater range in saturation, with the centroids of the color boxes at a greater distance from the origin. As with the measurements using the PR-715, there is overlap of the O2 and R3 data points. The increase in saturation may be due to the impact of sparkle (incomplete retroreflection) on the LabScan® XE measurements.[2] The mean color measurements for this color instrument separated by retroreflective sheeting type are shown in figure 29 in appendix B.

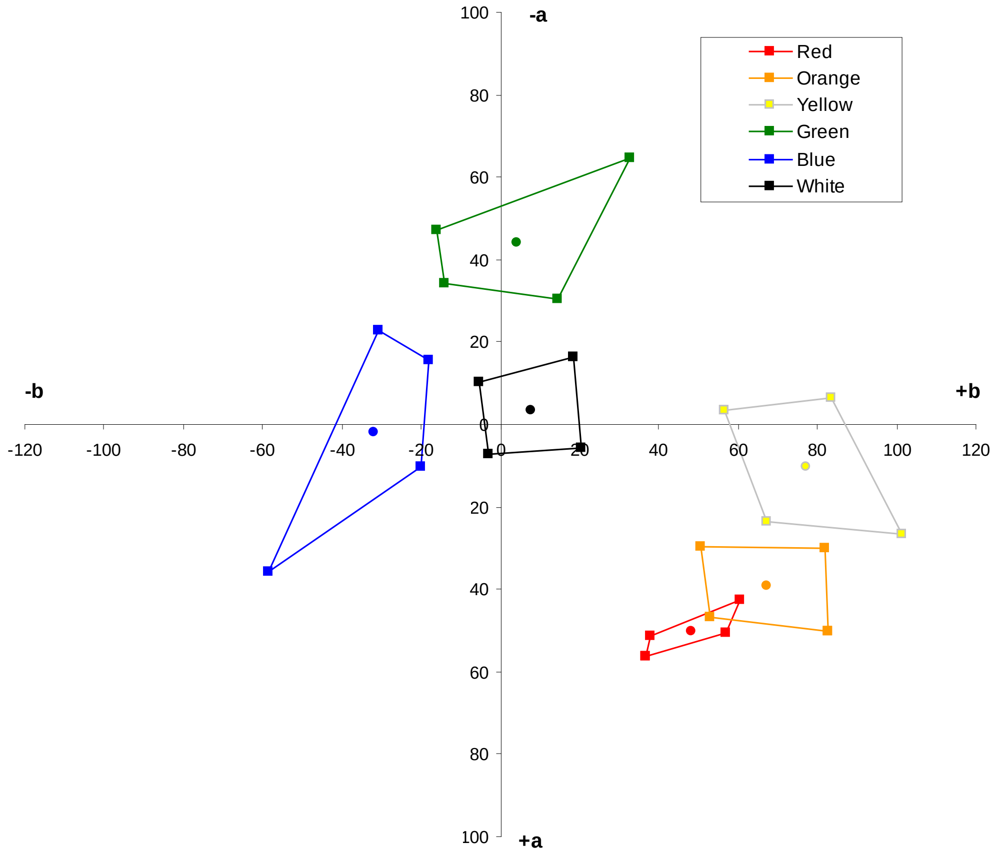


Figure 10. Graph. Laboratory (LabScan® XE) physical color measurements averaged over four retroreflective sheeting types.

Figure 11 and figure 12 show the physical measurements of the four individual retroreflective sheeting types that make up the averages depicted in figure 9. For the laboratory results, the measurements are tightly nested, with the type VIII material having slightly more saturated colors and the type III material having the least saturated colors (see figure 11). For the field results, the color boxes overlap rather tightly, except for red, orange, and yellow (see figure 12). For these three colors, the same pattern emerges in the relative saturation of the type VIII and type III materials but to a lesser degree. The type VIII material, which produced the most saturated colors in the laboratory, did not produce as saturated colors in the field.

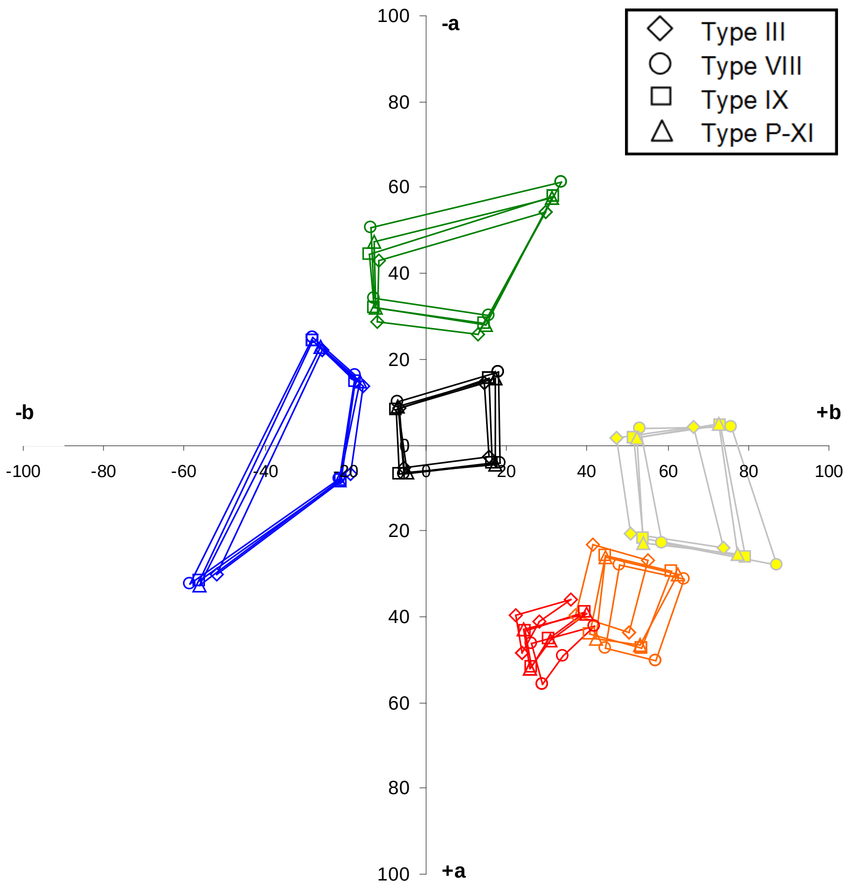


Figure 11. Graph. Laboratory (PR-715) physical color measurements of four retroreflective sheeting types.

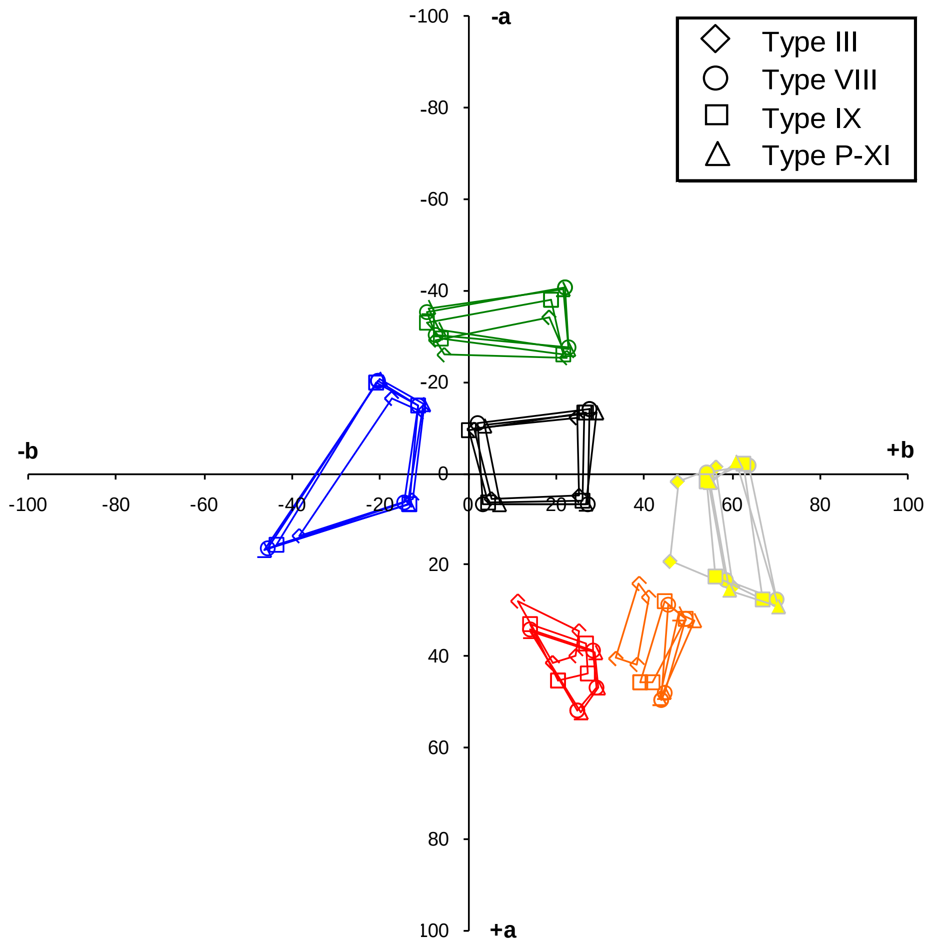


Figure 12. Graph. Field (PR-650) physical color measurements of four retroreflective sheeting types.

The variability of natural daylight conditions during the course of the study affected the luminance and CCT of the illuminant (the sky) and, thus, affected the measurements. A diffuse white standard reflector was measured before and after each block of color appearance trials to capture the variability of the daylight illumination conditions (e.g., passing clouds, high overcast, etc.) Figure 13 shows the entire set of field measurements for the diffuse white standard as measured under varying sky conditions throughout the study. As expected, the chromaticity coordinates varied due to the changing daylight illumination. Higher reflected luminance tended to be correlated with lower color temperature and lower reflected luminance with higher color temperature. This relationship is depicted in the legends for the extreme data points in figure 13. Under bright conditions, with abundant sunlight, the luminance is higher and the color temperature is lower, moving toward yellow. Passing clouds or high overcast blocked direct sunlight, resulting in reduced luminance and a higher color temperature. Thus, clear sunlight conditions tended to provide more spectral power in long wavelengths than did cloudy conditions. The measured data points fall fairly closely along the Planckian black body contour in the CIE chromaticity diagram, indicating an orderly progression of CCT similar to that of a black body radiator. These diffuse white measurements were used to adjust measurements for the individual color samples to account for the variable sky conditions during the course of a single day and across different days using the X, Y, Z to *L*a*b** transformation formulae.[4,10]

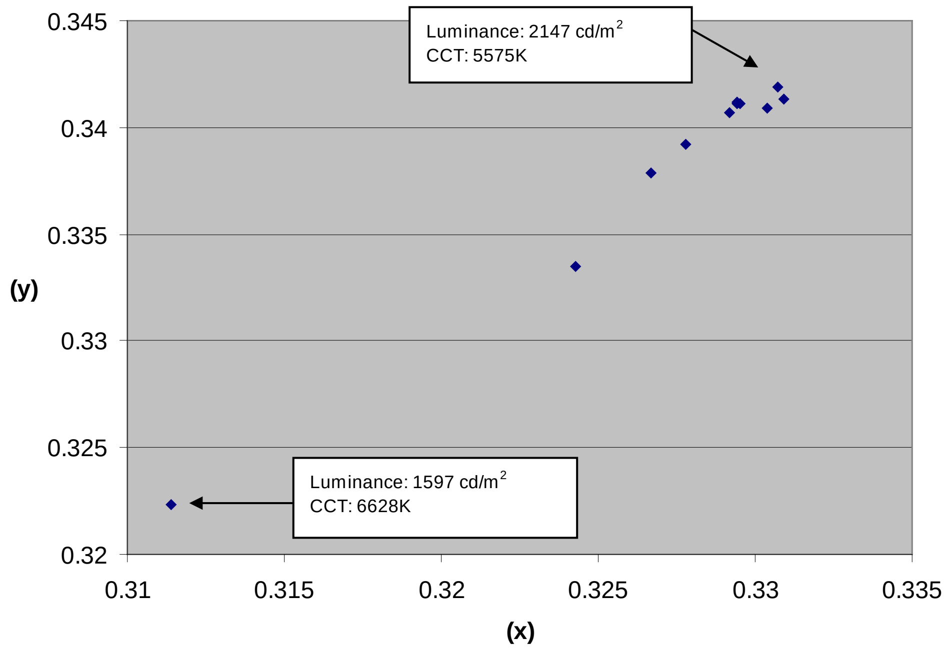


Figure 13. Graph. Individual field chromaticity measurements for the diffuse white standard reflector.

Perceptual Measurement Results

The human response data for hue and apparent saturation were plotted using UADs, which have two axes (B-Y and R-G) that range from -100 to +100. The B-Y axis represents the opponent colors of blue (-100) and yellow (+100). The R-G axis represents the opponent colors of red (-100) and green (+100). These diagrams represent a two-dimensional perceptual color space based on orthogonal red-green (R-G) and blue-yellow (B-Y) dimensions. The UADs are derived from the opponent process theory of human vision, where green and red operate as one antagonistic color pair, and blue and yellow operate as another antagonistic color pair. These two dimensions are similar to those used in the CIELAB formulation, with one axis (+*a*/-*a*) indicating the approximate red/green dimension and the other axis (-*b*/+*b*) indicating the approximate blue/yellow dimension. In the UAD formulation, a third dimension, designated *L*, provides an achromatic measurement from black to white. This third dimension is associated with the "lightness," or brightness, of the stimulus. For objective instrument measurements, the corresponding three dimensions (L, a, and b) can be combined into a three-dimensional color space, referred to as "LAB." CIELAB, a widely used version of this formulation, was used to describe the corresponding instrument measurements in this report.[(10)]

Comparison of Perceptual and Physical Measurements

Figure 14 shows the mean perceptual color ratings for the standard diffuse reflector with color filters for all 17 research participants. The labels in figure 14 identify the corners of the mean perceptual rating color boxes. The labels are the same as those used in figure 8, identifying the color (R, Y, O, G, B, and W) and filter number (1, 2, 3, and 4). Because of space considerations, the labels are only provided in this UAD. However, the corners of the color boxes stay in the same basic relative orientation for all the perceptual rating judgments (unless noted), and therefore, the labels in figure 14 can be referred to when viewing subsequent UADs. The mean perceptual measurements in figure 14 may be compared with the mean field measurements in figure 8. The scales for CIELAB and the UADs are not equivalent, so no absolute comparisons can be made; however, relative comparisons, such as shifts in opposite directions, can be made.

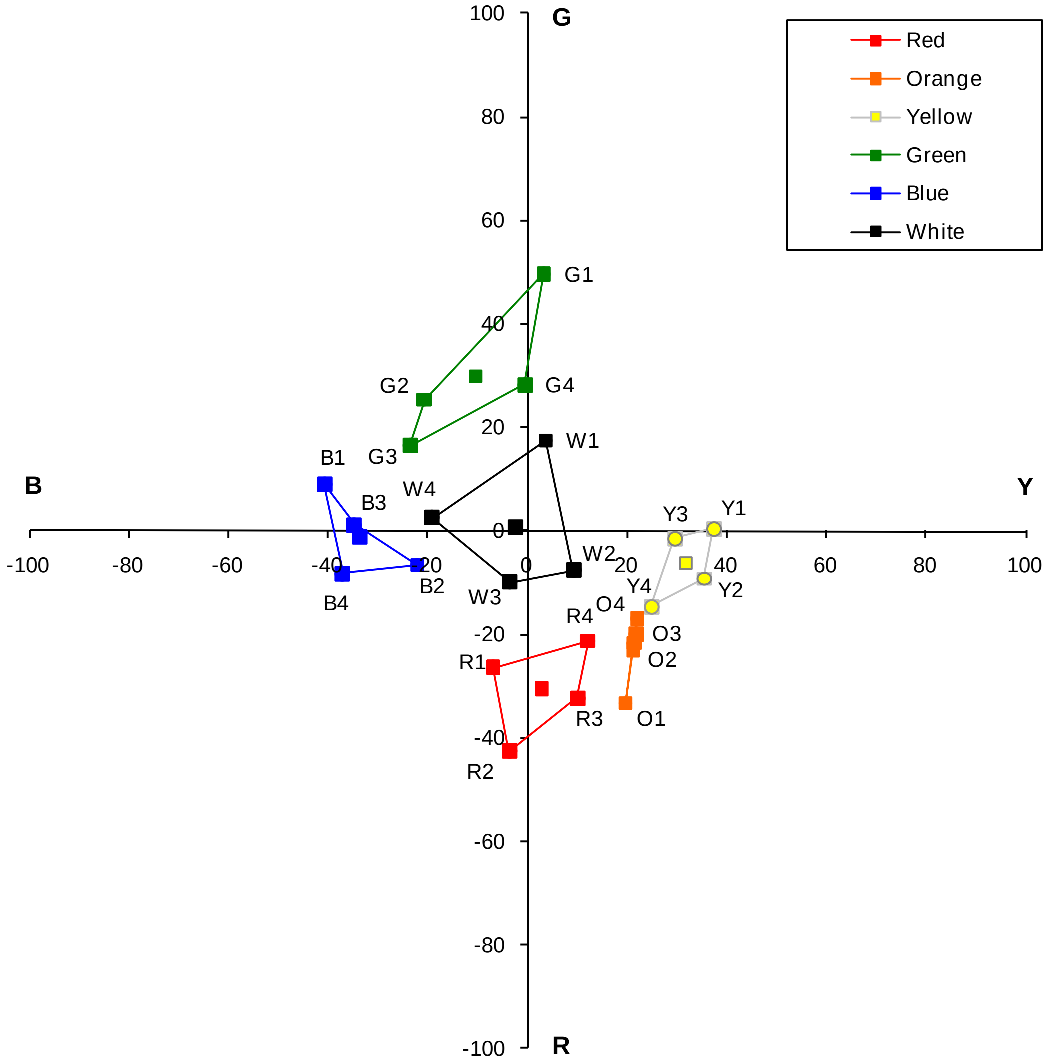


Figure 14. Graph. Mean perceptual color ratings for the white diffuse reflector with color filters for 17 participants.

A comparison between figure 8 and figure 14 reveals several important differences in the instrument and perceptual measurements of the reference colors. The most prominent of these differences relates to the red color box. The perceptual measurements show an expanded red color box that is not the smallest in size. For the perceptual measurements, the yellow box is somewhat compressed along the yellow dimension, and the orange box is extremely compressed along the yellow dimension. In the case of the orange box, this extreme degree of compression along the yellow axis indicates that, although instrument measurements show differing levels of saturation along the +*b* axis, human observers show an insensitivity to these differences. The human observers respond with the same amount of yellow in their color ratings for all four corners of the orange box. The red, orange, and yellow color boxes are closest in terms of hue angle. In figure 14, the constant hue lines separating the orange and yellow boxes reveal no direct overlap, but the hue angle line separating the red and orange boxes indicates an overlapping of these two color designations. Some degree of perceptual color confusion might be expected in this area. The perceptual results for the white, green, and blue color boxes show distinct color separations.

The mean perceptual color measurements for all 17 participants averaged over the 4 retroreflective material types are shown in figure 15. The standard errors for the means along either dimension (R-G or B-Y) ranged from 0.61 to 0.62; so mean differences greater than 1.2 scale units are likely to be statistically significant. The mean perceptual color measurements in figure 15 may be compared to the mean laboratory color measurements for the same stimuli in figure 9. Since the scales are not equivalent, only certain relative comparisons may be made. Such a comparison of figure 9 and figure 15 for the retroreflective samples revealed many of the same basic differences found in the comparison of figure 8 and figure 14 for the diffuse color samples. The green and white boxes both changed from the yellow side to the blue side of the neutral axis. For perceptual measurements, the yellow color box and orange color box are compressed along the saturation dimension, with the orange box more compressed than the yellow box.

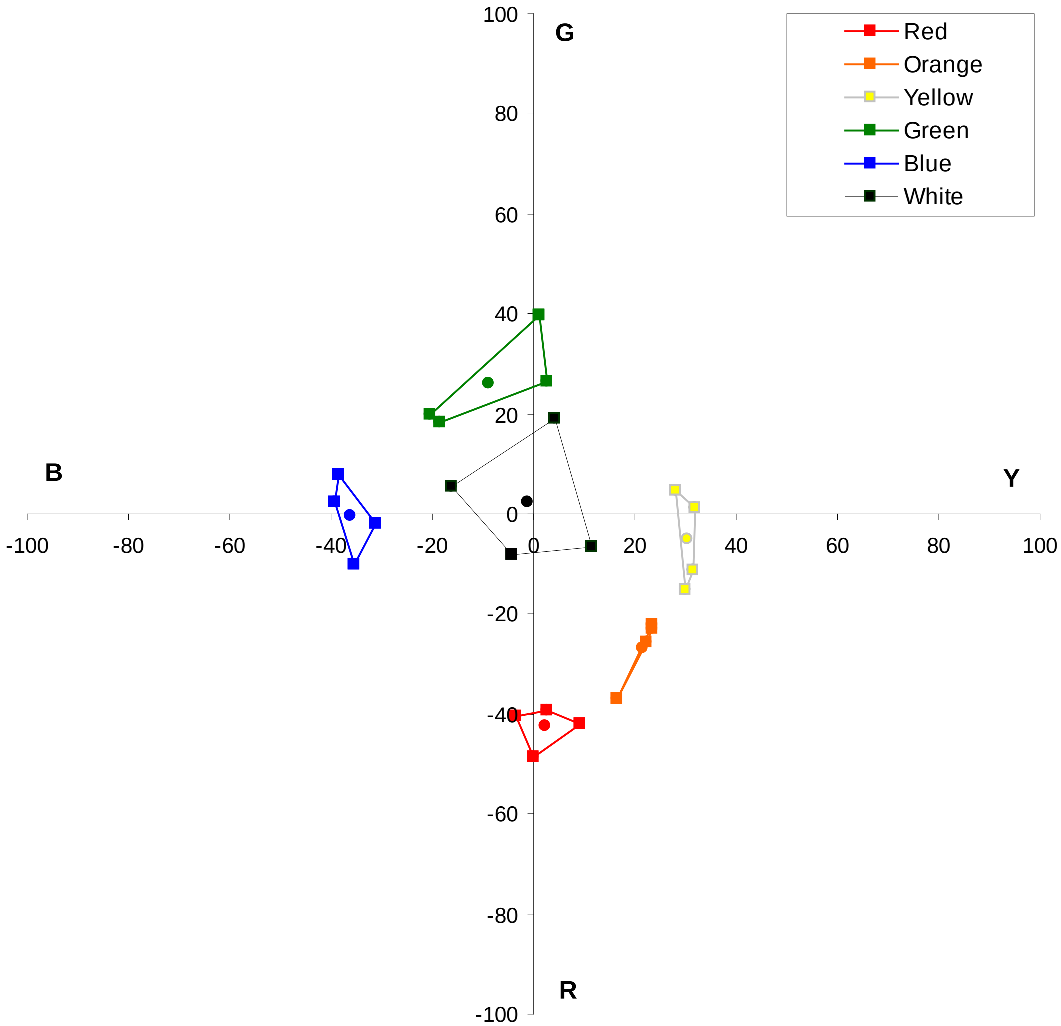


Figure 15. Graph. Mean perceptual color ratings averaged over four retroreflective sheeting types for 17 participants.

Comparison of Perceptual Measurements for Diffuse versus Retroreflective Samples

In addition to the data provided in figure 15, the mean color ratings given by each participant for the type VIII material are in appendix C to provide an indication of the between-participant variability inherent in the mean measurements. Figure 15 may be compared with figure 14 to reveal differences in the shapes and locations of the perceived color boxes for retroreflective materials relative to the diffuse color samples. The most prominent difference is that the red color box in figure 15 has moved significantly toward a more saturated red but its size has been substantially reduced. The yellow color box has become even more compressed along the yellow dimension. The orange box remains extremely compressed, although its orientation has changed slightly. For the retroreflective materials, the orange box is not oriented perpendicular to the B-Y axis, indicating slightly greater sensitivity to differences in yellow content. The red, orange, and yellow color boxes remain closest in terms of hue angle, but there are no direct overlaps of the mean perceptual measurements. The results for the white, green, and blue color boxes continue to show distinct color separations.

Hue Line Separation of Colors and Saturation

In figure 15, hue angle lines between the orange and yellow boxes and the red and orange boxes reveal clear hue separations between the colors. This outcome may indicate adequate perceptual discrimination among these colors despite the fact that some of the instrument measurements overlapped. A comparison of the perceptual measurements of diffuse color boxes in figure 14 and retroreflective color boxes in figure 15 does not demonstrate the same reduction in saturation for the diffuse colors as was indicated with the instrument measurements (see figure 8 and figure 9). This may indicate that human observers are not as sensitive to reductions in saturation or that the mechanism that resulted in reduced saturation for the diffuse color samples was related to the instruments used. The perceptual measurements resulted in lower overall saturation levels than the instrument measurements. This change toward the lower end of the saturation scale may also have resulted in a reduced sensitivity to changes in saturation. Finally, the tendency of human observers to attempt to maintain color constancy over a variety of different lighting and reflection variables may account for this apparent reduction in sensitivity.

Differences in the perceptually measured mean color boxes for the four individual retroreflective sheeting types are shown in figure 16. The figure reveals numerous data points close to the hue angle lines separating yellow and orange as well as orange and red. This proximity indicates the potential for perceptual color confusion. These areas, which are similar to the border regions of the FHWA color boxes, are of particular interest and were the basis for conducting further analyses.(3) The specific points of interest were comparisons of O1 to R4, O2 to R3, Y3 to O3, and Y4 to O4.

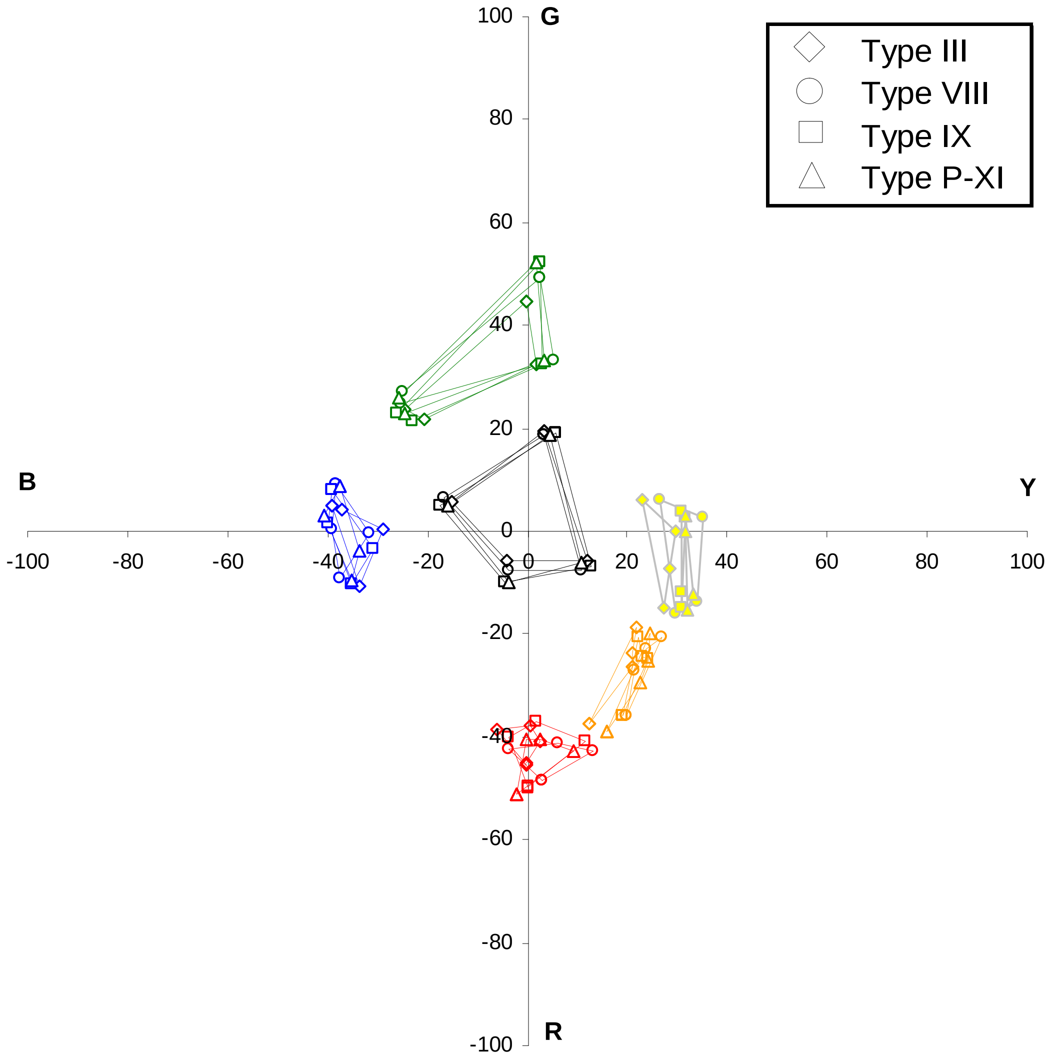


Figure 16. Graph. Mean perceptual color ratings of four sheeting types.

The data points of interest were converted from a Cartesian to a polar coordinate system. This method has been used to measure differences in hue angles in the CIELAB color space and was applied here to measure differences in hue angles in UADs.(14) The means of the hue angles for data points of interest were plotted with their respective 95 percent confidence limit (two standard errors of the mean) to determine potential color/sheeting combinations for comparison. Figure 17 shows the mean hue angle with two standard errors of the mean for red (corners 3 and 4), orange (corners 1, 2, 3, and 4), and yellow (corners 3 and 4) for all four retroreflective sheeting types.

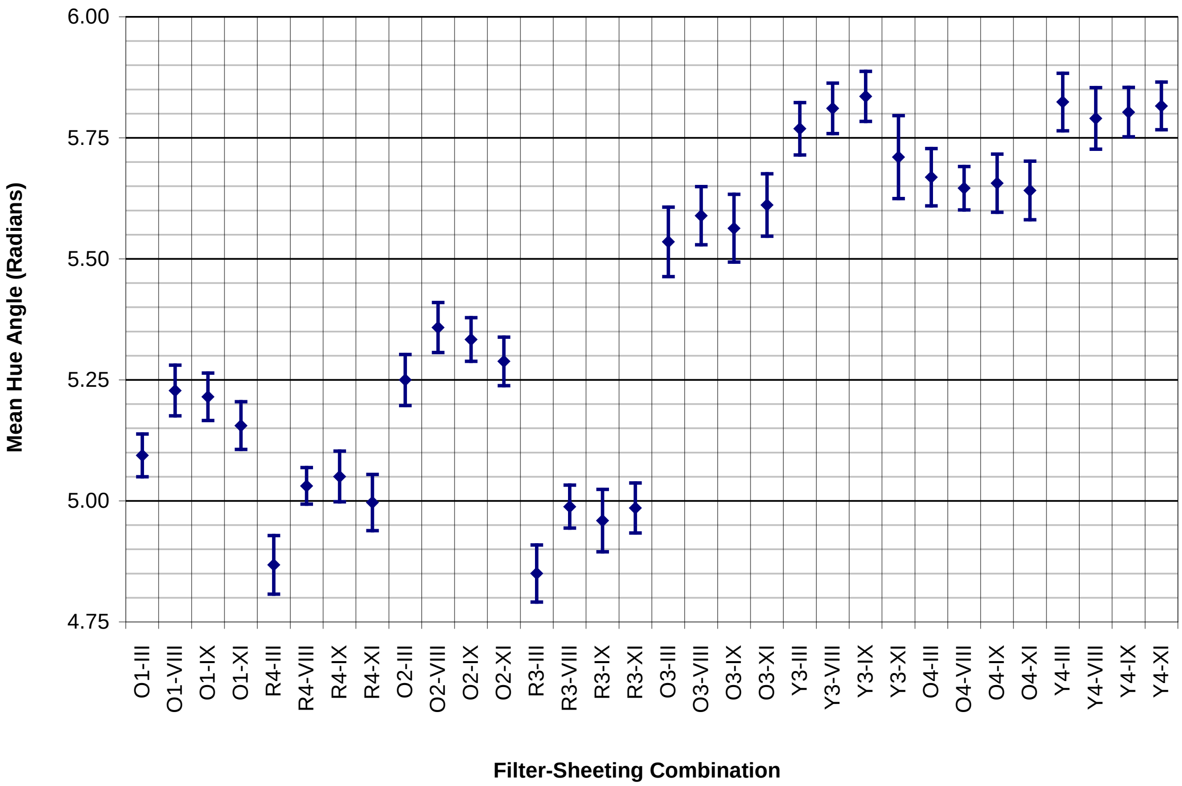


Figure 17. Graph. Mean hue angle in radians with 95 percent confidence limit of two standard errors.

Those data points in figure 17 that are located within the 95 percent confidence limit of another data point were evaluated using paired comparison *t*-tests. There were 13 such pairs of data points. To control the familywise error rate due to multiple *t*-tests, a Bonferroni post-hoc adjustment was made so that the probability required at the 0.05 significance level would be 0.004 ($p = 0.05/13$). The results in table 3 indicate that observed differences were not statistically significant in 11 out of 13 color/sheeting combinations. These 11 pairs were the O1 type III and R4 type VIII, O1 type III and R4 type IX, O1 type III and R4 type P-XI, O3 type P-XI and Y3 type P-XI, O3 type VIII and Y3 type P-XI, O3 type IX and Y3 type P-XI, O4 type III and Y4 type VIII, O4 type III and Y3 type P-XI, O4 type VIII and Y3 type P-XI, O4 type IX and Y3 type P-XI, and O4 type P-XI and Y3 type P-XI paired comparisons. For these paired comparisons, non-significant differences indicate no statistical difference, and therefore, some degree of perceptual color confusion might be expected between signs having any of these 11 unique color and sheeting combinations. Specifically, the comparisons between O1 type III and R4 type IX and between O4 type III and Y3 type P-XI had the smallest mean hue angle difference, 0.041 radians, and might be expected to produce the most color confusion. In table 3, the asterisk means that the difference was statistically significant and, therefore, of less interest.

Table 3. Paired comparison *t*-tests for hue angles.

Paired Comparison	Paired Differences: Mean Hue Angle Difference (Radians)	Standard Deviation	Standard Error Mean	95 Percent Confidence Interval of the Difference: Upper	Lower	*t*	Degrees of Freedom	Significance (2-tailed)
Orange 1 type III, Red 4 type VIII	0.063	0.099	0.024	0.012	0.114	2.612	16	0.019
Orange 1 type III, Red 4 type IX	0.041	0.137	0.033	-0.029	0.112	1.251	16	0.229
Orange 1 type III, Red 4 type P-XI	0.094	0.143	0.035	0.020	0.167	2.698	16	0.016
Orange 1 type P-XI, Red 4 type IX	0.103	0.118	0.029	0.042	0.164	3.587	16	0.003*
Orange 3 type P-XI, Yellow 3 type P-XI	-0.101	0.174	0.042	-0.191	-0.011	-2.385	16	0.030
Orange 3 type VIII, Yellow 3 type P-XI	-0.121	0.163	0.039	-0.205	-0.037	-3.068	16	0.007
Orange 3 type IX, Yellow 3 type P-XI	-0.147	0.214	0.052	-0.257	-0.037	-2.823	16	0.012
Orange 4 type III, Yellow 4 type VIII	-0.122	0.150	0.036	-0.199	-0.044	-3.344	16	0.004
Orange 4 type IX, Yellow 4 type VIII	-0.134	0.125	0.030	-0.198	-0.069	-4.402	16	0.000*
Orange 4 type III, Yellow 3 type P-XI	-0.041	0.181	0.044	-0.134	0.051	-0.946	16	0.358
Orange 4 type VIII, Yellow 3 type P-XI	-0.064	0.171	0.042	-0.152	0.024	-1.539	16	0.143
Orange 4 type IX, Yellow 3 type P-XI	-0.054	0.196	0.048	-0.155	0.047	-1.128	16	0.276
Orange 4 type P-XI, Yellow 3 type P-XI	-0.069	0.148	0.036	-0.145	0.007	-1.916	16	0.073

* Indicates statistically significant at adjusted 0.05 level (> 0.004).

Summary of Measurement Results

In summary, laboratory physical color measurements of the white diffuse reflector with a color filter showed well-defined color boxes arranged in an orderly manner when expressed in terms of CIELAB plots. The only major unexpected outcome was the relatively small size of the red color box. In the case of retroreflective colors, the laboratory measurements revealed that retroreflective properties tend to reduce the saturation of all colors. The resulting smaller separation between some of the color boxes may make them more difficult to discriminate. The reduced size and closer proximity of the red, orange, and yellow color boxes for retroreflective materials indicate a higher probability of color confusion between certain orange and yellow signs and between certain orange and red signs. Perceptual color measurements revealed an additional consideration. For retroreflective materials, the orange color box collapsed and the yellow color box became extremely compressed along the yellow axis of a UAD. This compression along the yellow dimension indicates that the human observers were relatively insensitive to saturation changes along the yellow dimension for yellow and orange sign materials. Nevertheless, constant hue lines for perceptual measurements showed a greater separation between orange and yellow and between orange and red than in the physical measurements. The results of the study indicate that human observers are able to discriminate colors in this region of the spectrum better than might be inferred from CIELAB plots of physical measurements. The perceptual results for the white, green, and blue color boxes showed distinct color separations, as did the physical measurements.

LUMINANCE AND BRIGHTNESS

The human response data for brightness were collected using two different methodologies. The brightness rating data were collected for each color sample presented in the color appearance portion of the experiment. Participants rated each color sample on a scale of 0–100, with 0 being the darkest and 100 being the brightest. The brightness ranking data were collected for two subsets of color samples (yellow and red). Participants ranked each of the five color samples in descending order from 1 to 5, with 1 being the brightest and 5 being the darkest.

Luminance

Figure 18 shows the mean luminance of the color samples measured by the PR-650 during the color appearance portion of the study. The mean values are averaged across seven replications of the same color area corner. The error bars represent ±2 standard errors of the mean (95 percent confidence limit).

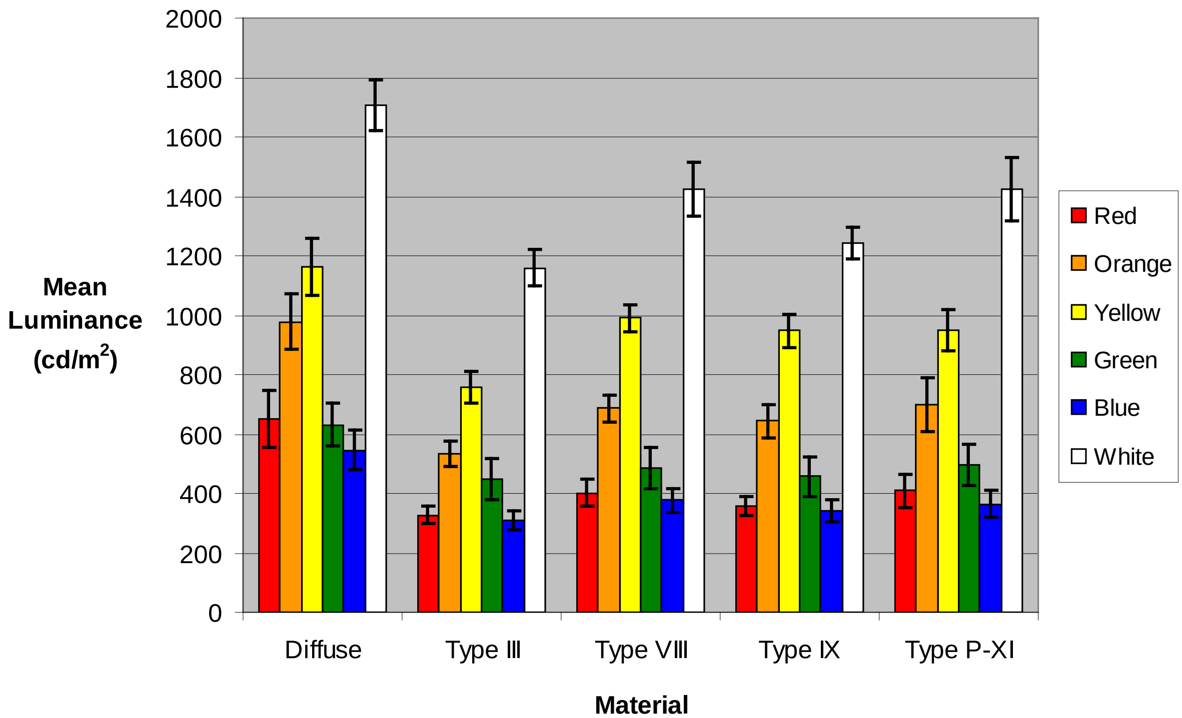


Figure 18. Graph. Mean field luminance measurements by sheeting type for all colors.

The five colors in figure 18 generally followed a strict order of relative luminance, with a peak in orange and yellow and lower luminance values for red and blue, the extremes of the visibility spectrum. White had a consistently higher mean luminance than any other color within each sheeting type; however, in general, the remaining colors retained their same relative position across the sheeting types. A comparison of the white samples reveals that type III sheeting had the lowest mean luminance. The other three retroreflective types had similar luminance values, and the diffuse white sample had the highest measured mean luminance. The diffuse color samples had the highest mean luminance values for each color when compared to the respective mean luminance values of the retroreflective samples, and the type III samples had the lowest luminance values. The relatively small error bands indicate that most of these luminance differences are likely to be statistically significant.

Figure 19 shows the mean luminance values of the Y2 and R2 color samples. These were the only color samples used in the ranking portion of the color appearance determinations. The error bars represent ±2 standard errors of the mean. The relatively high variability is not surprising, given the small number of observations and the variable number of measurements for each color sample. Figure 19 illustrates that the diffuse color samples had the highest mean luminance for both yellow and red, as evidenced by the lack of overlapping error bars. The mean luminance values for the four retroreflective sheeting types for each color were similar, with no apparent differences. Overall, the red colors produced far lower luminance values than the yellow colors. A comparison of the relationships in figure 18 and figure 19 reveals similar patterns. This similarity is to be expected, given that the samples used in figure 19 represent a subset of the samples used in figure 18.

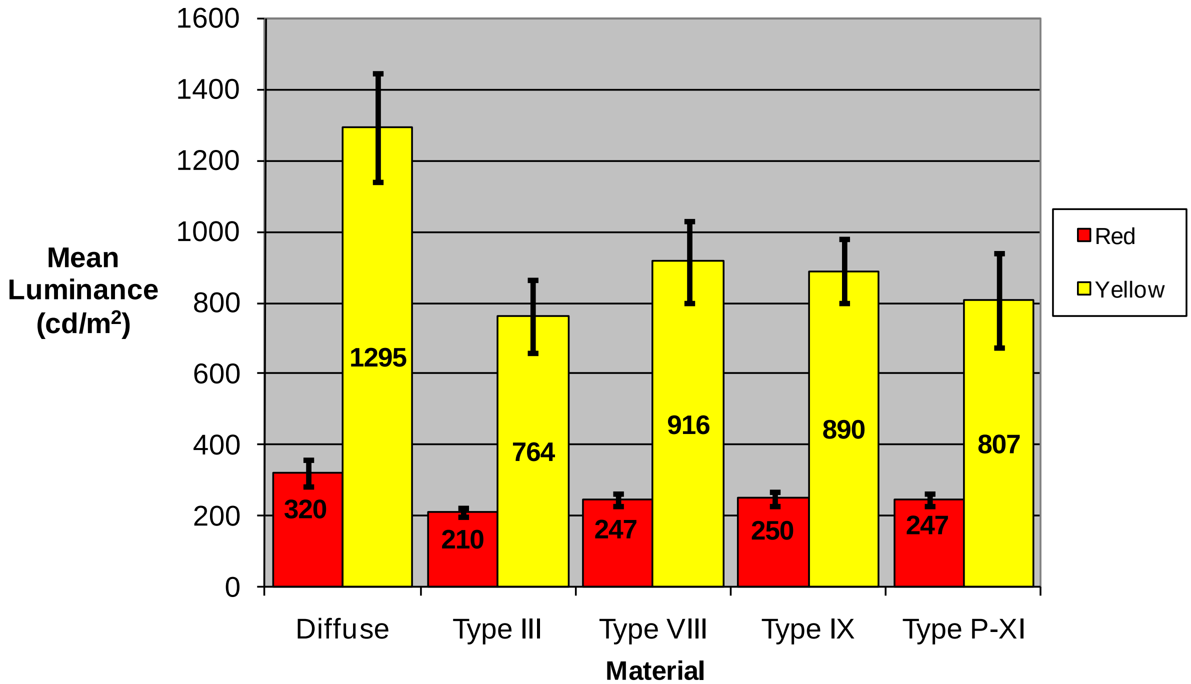


Figure 19. Graph. Mean field luminance measurements of the yellow and red samples used for the brightness ranking task.

Brightness Ratings

Figure 20 shows the mean brightness ratings by sheeting type for all colors from the perceptual ratings. In this case, the means are averaged across all four color areas, across all replications of measurements, and across all 17 research participants. The error bars represent ±2 standard errors of the mean (95 percent confidence limit).

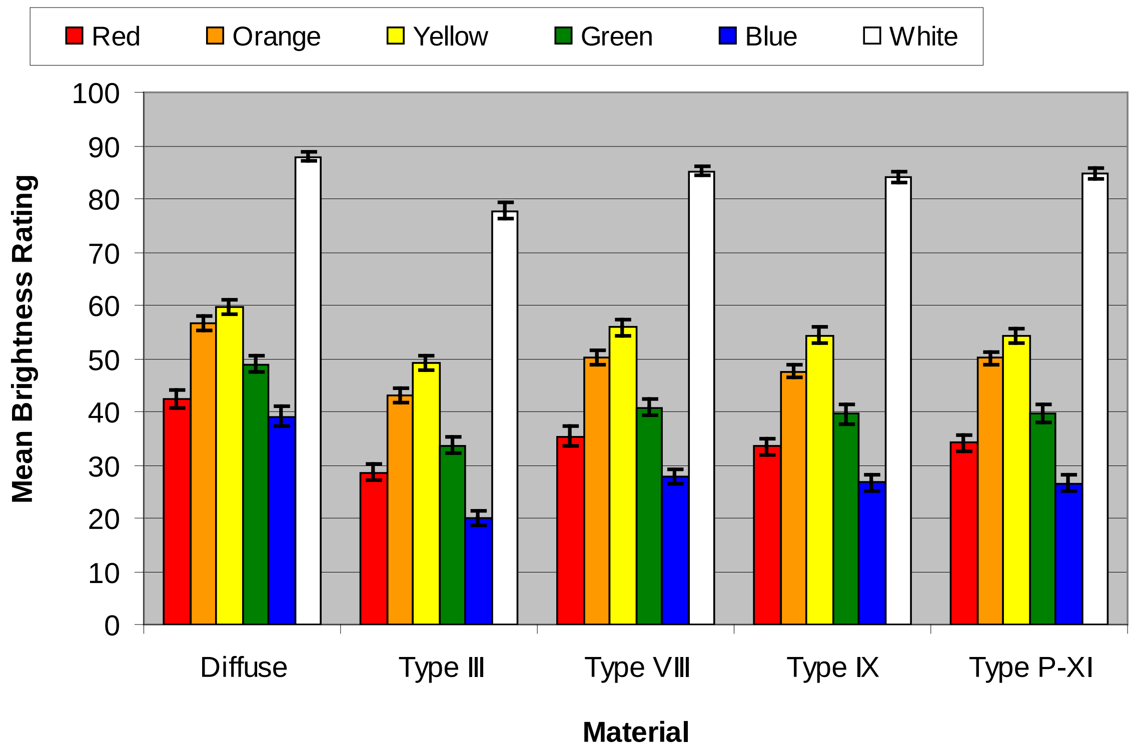


Figure 20. Graph. Mean brightness ratings by sheeting type for all colors.

Figure 20 shows the effect of color and sheeting on human brightness ratings. Based on the measured field luminance values in figure 18, it was anticipated that the diffuse samples would have higher ratings and type III sheeting would have lower ratings than types VIII, IX, and P-XI. This outcome is shown in figure 20, where the diffuse reflector is generally the brightest and the type III sheeting is generally the least bright for a given color. Consequently, no further analysis was done regarding these two materials. However, it was unclear how sheeting types VIII, IX, and P-XI would compare among each other, and so additional analyses were conducted on these materials. Within a given color, post-hoc paired-sample *t*-tests were conducted to determine if there were any statistically significant differences between certain pairs of sheeting types. The color-sheeting combinations chosen for paired comparisons were the types VIII, IX, and P-XI with overlapping 95 percent confidence ranges (two standard errors). Based on the results of the initial analysis, all three sheeting types for each color were compared, which resulted in 18 post-hoc paired comparison tests. To control the familywise error rate due to multiple *t*-tests, a Bonferroni post-hoc adjustment was made so that the probability required at the 0.05 significance level would be 0.003 (p = 0.05/18). Table 4 shows the results of the selected *t*-tests, of which only one was statistically different at the Bonferroni-adjusted significance of 0.003. In all of the other 17 cases, the observed differences in brightness ratings were not statistically significant. Brightness discrimination among these comparisons would not be expected to be very good. Furthermore, brightness discrimination of this sort is of less practical importance in driving situations. Overall the results indicate that the type P-XI sheeting is not different in terms of the human daytime brightness rating response than the two other microprismatic sheeting types tested (types VIII and IX).

Table 4. Paired comparison *t*-tests for brightness ratings.

Paired Comparison	Paired Differences: Mean Rating Difference (percent)	Paired Differences: Standard Deviation	Paired Differences: Standard Error Mean	Paired Differences: 95 Percent Confidence Interval of the Difference: Upper	Paired Differences: 95 Percent Confidence Interval of the Difference: Lower	*t*	Degrees of Freedom	Significance (2-tailed)
Blue type VIII, Blue type IX	1.12	2.68	0.65	-0.26	2.50	1.73	16	0.104
Blue type VIII, Blue type P-XI	1.197	3.316	0.804	-0.508	2.901	1.488	16	0.156
Blue type IX, Blue type P-XI	0.074	2.883	0.699	-1.408	1.556	0.106	16	0.917
Green type VIII, Green type IX	1.256	3.476	0.843	-0.531	3.043	1.490	16	0.156
Green type VIII, Green type P-XI	1.186	2.676	0.649	-0.190	2.562	1.827	16	0.086
Green type IX, Green type P-XI	-0.070	2.974	0.721	-1.600	1.459	-0.098	16	0.923
Orange type VIII, Orange type IX	2.537	3.111	0.755	0.937	4.137	3.362	16	0.004
Orange type VIII, Orange type P-XI	0.074	2.304	0.559	-1.111	1.259	0.132	16	0.896
Orange type IX, Orange type P-XI	-2.463	2.694	0.653	-3.848	-1.078	-3.770	16	0.002*
Red type VIII, Red type IX	1.894	2.869	0.696	0.419	3.369	2.722	16	0.015
Red type VIII, Red type P-XI	1.226	2.950	0.715	-0.291	2.742	1.713	16	0.106
Red type IX, Red type P-XI	-0.669	2.613	0.634	-2.012	0.675	-1.055	16	0.307
White type VIII, White type IX	1.109	1.743	0.423	0.212	2.005	2.622	16	0.019
White type VIII, White type P-XI	0.487	1.574	0.382	-0.322	1.296	1.276	16	0.22
White type IX, White type P-XI	-0.621	1.654	0.401	-1.472	0.229	-1.549	16	0.141
Yellow type VIII, Yellow type IX	1.434	3.076	0.746	-0.148	3.015	1.922	16	0.073
Yellow type VIII, Yellow type P-XI	1.488	2.669	0.647	0.116	2.860	2.299	16	0.035
Yellow type IX, Yellow type P-XI	0.054	2.676	0.649	-1.322	1.430	0.083	16	0.935

* Indicates statistically significant at adjusted 0.05 level (> 0.003).

The preselection process for the paired comparison *t*-tests eliminated type III sheeting, which was determined to be significantly different from the other types of sheeting in terms of brightness ratings. This determination is supported by figure 20. For any given color, the brightness rating of the type III material was consistently lower than the rating for any of the other materials.

In general, the mean brightness ratings in figure 20 show similar patterns to the mean field luminance measurements in figure 18. They both reveal the same general shape among the five colors in strict order from lowest to highest: blue, red, green, orange, and yellow. For the white color, type III sheeting had the lowest mean brightness rating followed by the other three retroreflective types, which all had similar intermediate ratings, and the diffuse color samples, which had the highest brightness rating. A high degree of similarity was observed between the field instrument measurements of mean luminance in figure 18 and the mean brightness ratings in figure 20.

Brightness Rankings

Figure 21 shows the mean brightness rankings for the Y2 and R2 samples. The brightness rankings have been inverted to be comparable to the brightness ratings in figure 19. In the actual testing, a rank of 1 represented the brightest sample and a rank of 5 represented the darkest sample. The diffuse samples had the highest mean brightness (lowest mean rank), and type III sheeting had the lowest mean brightness (highest mean rank) for both colors. Sheeting types VIII, IX, and P-XI tended to represent an intermediate brightness. These relationships are similar to those found for the mean field luminance measurements for this same subset of stimuli and for the mean brightness ratings given by the participants (see figure 19 and figure 20). Thus, field luminance measurements are predictive of both brightness ratings and brightness rankings, which were in themselves similar. In short, the two perceptual response modes tended to complement each other, and both could be predicted from physical measurements of stimulus luminance. The small error bands in figure 21 reveal the overall consistency in the observed brightness ranking judgments. Since the samples were placed side-by-side in the brightness rankings, the rankings might be expected to show more sensitivity to differences in luminance than the brightness ratings, where the samples were presented one at a time. The mean luminance values for each of the samples are given in figure 21, as well as the mean rankings. For red, the mean rankings track the mean luminance values closely. For yellow, the same close tracking is present, with a small deviation for type IX, which has a mean ranking somewhat lower than might be expected based on the mean luminance.

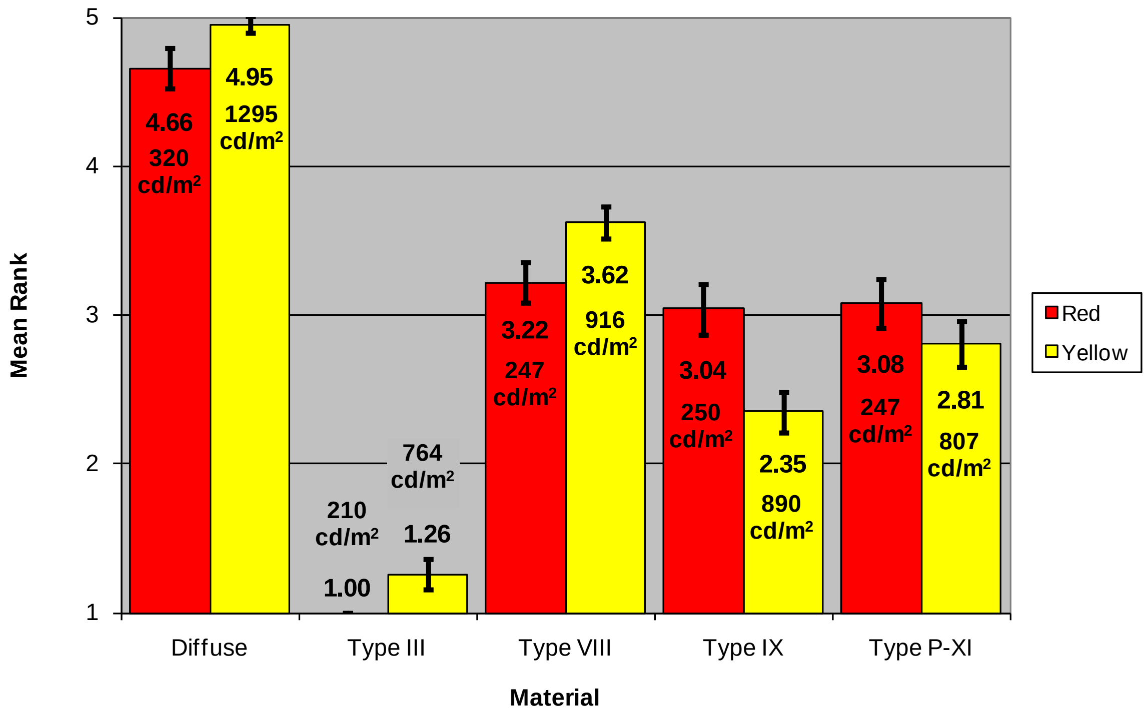


Figure 21. Graph. Mean brightness rankings for yellow and red samples.

In terms of the effects of retroreflective sheeting material, both the brightness rating and ranking methods revealed that the diffuse color samples had the highest brightness. Instrument measurements also revealed this condition to have the highest luminance. The presence of any retroreflective property decreased the overall brightness of the color sample, with encapsulated glass beads (type III) having the strongest effect on reducing brightness. Instrument measurements of luminance generally confirmed these relationships.

Brightness as a Function of Luminance

Figure 22 shows the mean brightness ratings portrayed in figure 20 as a function of the mean luminance by type of material for each of the six colors tested. Linear regression lines are drawn for the quadrant samples constituting a given color area. In general, the colors with higher mean luminance values had higher mean brightness ratings. Thus, across colors, the brightness rating of a particular color could be predicted to some degree by the overall luminance of the color, as previously described. However, there was a substantially weaker relationship between brightness and luminance within a given color, except in blue and green. For all other colors, large changes in luminance produced only small changes in brightness within the color. The slopes of the curves for each of these four remaining colors were relatively shallow and approximately equal in magnitude. The relative strengths of the linear regression fits within each color are given by the R^2 values at the right side of the figure. The general shape of the function across colors illustrates the expected logarithmic relationship between perceptual brightness and physical luminance. Thus, a logarithmic function was fit to all of the data for all five colors plus white and is shown as a dashed line.

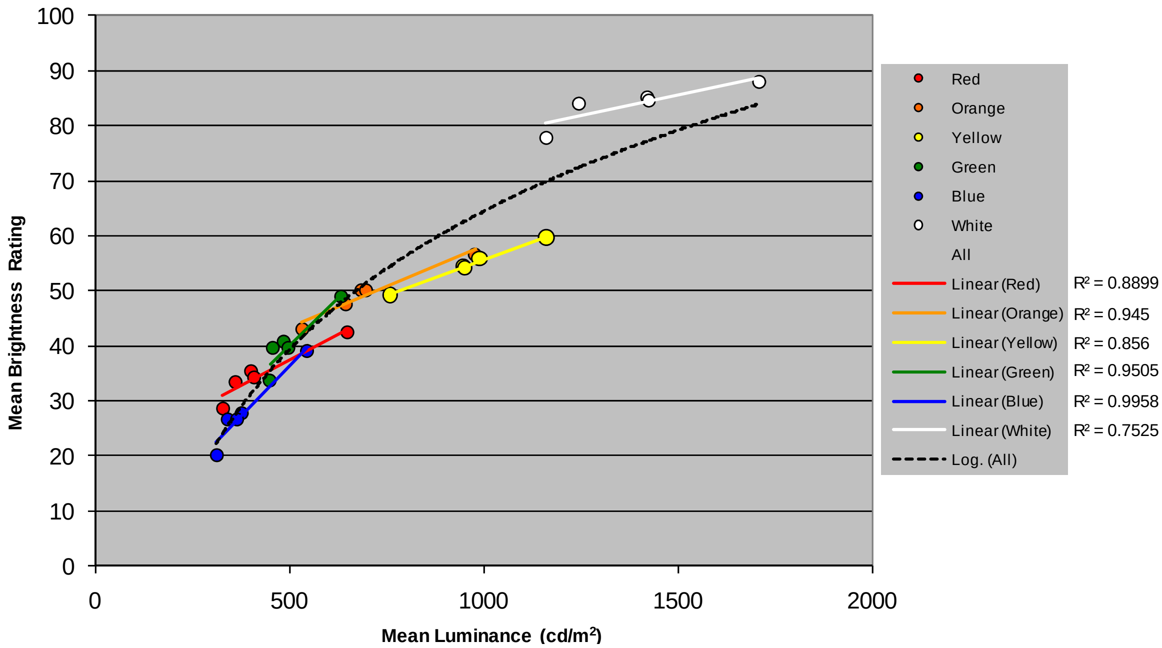


Figure 22. Graph. Mean brightness rating as a function of mean luminance for six colors.

Figure 23 shows the mean brightness ratings portrayed in figure 20 as a function of the mean lightness (*L**) of each sample for each of the six colors tested. The dashed line represents a linear function fit to all of the data for all five colors plus white. Comparing the dashed lines in figure 22 and figure 23, figure 22 shows the anticipated logarithmic function for a physical measure of the stimulus luminance, and figure 23 shows the anticipated linear function for a subjective measure of the stimulus lightness (brightness).

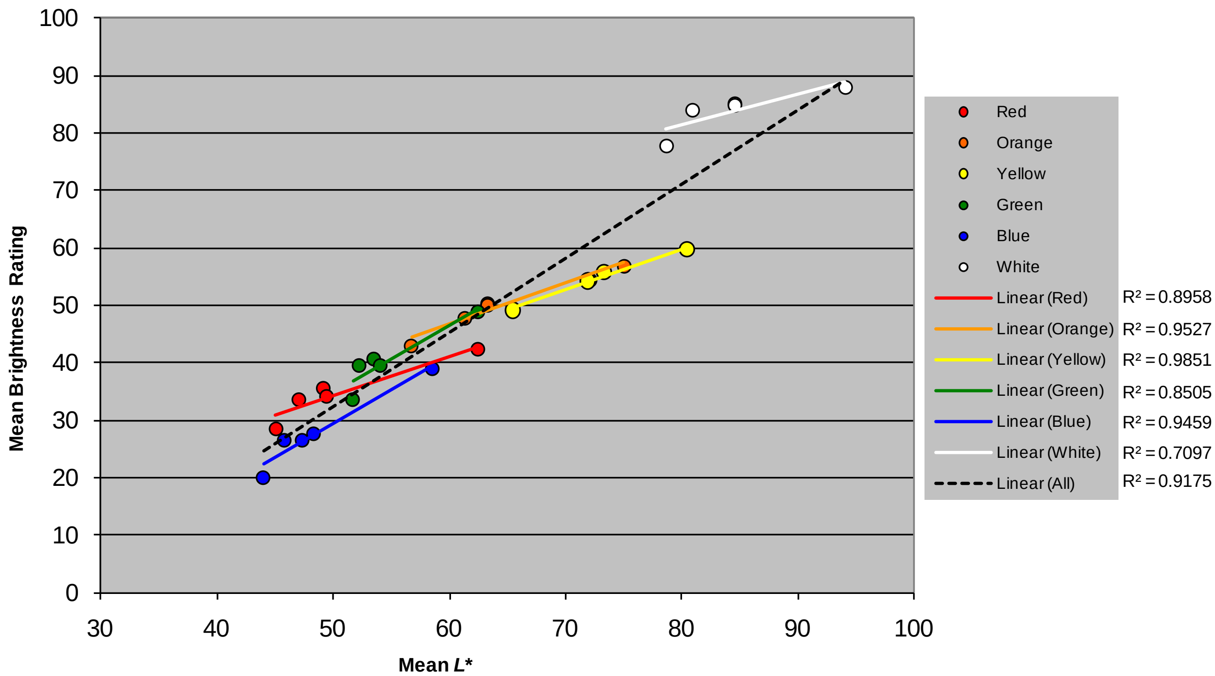


Figure 23. Graph. Mean brightness rating as a function of mean lightness (*L) for six colors.**

In summary, as concerns perceptual measures of brightness, the diffuse samples were always judged the brightest for any given color. Retroreflective properties tend to make colors appear darker relative to this reference condition, and the use of encapsulated glass beads (type III) had the greatest dimming effect. Instrument measurements of luminance generally confirmed these relationships. Thus, perceptual brightness relationships can be predicted by physical measurements of luminance. The overall reduction in luminance that accompanies retroreflective materials may partially explain the loss of saturation of these colors when compared to the diffuse reference condition. As the luminance of the sample is reduced, chromatic saturation is also reduced.

CHAPTER 4. SUMMARY AND DISCUSSION

The effects of retroreflective properties on the chromaticity and luminance of reflected light under natural daylight and simulated daylight (D65) conditions were explored. Chromaticity and luminance values were also compared with the daytime appearance of the materials as judged by a group of human observers. Understanding the effects of retroreflective properties on physical measurements and perceptual ratings is important to FHWA in defining the size and shape of the color areas used to specify colors for traffic control signs. The color of traffic control signs is an important part of the coding system for conveying consistent information to drivers. On one hand, if the color areas are too large, some of them may come close to functionally overlapping, especially when the variability of human color judgments is taken into account, resulting in possible color confusion and misidentification by drivers. On the other hand, if the color areas are too small, manufacturers can be unnecessarily restricted in the types of materials and processes used for the production of street and highway signs, possibly reducing alternatives and increasing costs. A balance must be struck between these and other competing concerns in defining the size and shape of the FHWA color areas. Information on how drivers perceive the colors of traffic control signs is important in this regard. Equally important is information on how to objectively measure these colors so manufacturers can efficiently and reliably meet FHWA color requirements.

In the present experiment, a white diffuse reflector coupled with appropriate color and neutral density filters was employed as a reference standard. CIELAB plots representing laboratory physical color measurements of the reference standard show well-defined color areas that are arranged in an orderly manner in the color space. Based on the sizes and orientations of the color areas in terms of chromaticity coordinates (see figure 2), the only major unexpected outcome in the physical measurements is the relatively small size of the red color area. The laboratory physical measurements show that retroreflective materials tend to reduce the saturation of all of the colors relative to the diffuse reference. Based solely on physical measurements, the reduced size and closer proximity of the red, orange, and yellow color areas for retroreflective materials might indicate a higher probability of color confusion between certain orange and yellow signs and between certain orange and red signs.

Perceptual color measurements, expressed in terms of UADs, also show reduced saturation for retroreflective materials relative to the diffuse reference. This might be expected, since retroreflective properties reduced the luminance of the samples. This outcome confirms earlier results found by Davis and Miller.[(9)] The present experiment revealed an additional consideration; for retroreflective materials, there was significant insensitivity in the perception of differences in yellow saturation for the orange and yellow color boxes. This compression along the yellow dimension indicates that the human observers were relatively insensitive to chromaticity changes along the yellow dimension for retroreflective materials. This relative human insensitivity could have implications for the weathering or fading of yellow traffic control signs. Yellow signs usually fade toward white, representing a reduction in saturation along the yellow color dimension. While not desirable, the fading of yellow sign materials may not be as noticeable to drivers as might be anticipated from instrument measurements. In any case, constant hue lines for perceptual measurements show a clearer separation between orange and yellow and between orange and red than the instrument measurements indicated. To a certain degree, human observers seem to be able to discriminate between different colors in this region of the spectrum better than might be inferred from physical measurements. The results for the white, green, and blue color areas show distinct color separations for all retroreflective sign materials.

Overall, the results of the experiment revealed that, under daytime lighting conditions, drivers with normal color vision should not experience confusion discriminating hue among the six FHWA color boxes that were evaluated. On average, across all retroreflective sheeting types, the separation among the perceptual color areas appeared sufficient to discriminate between any two colors of the six tested. However, when individual sheeting types were considered, there were two cases where the differences between combinations of color and sheeting type were not statistically significant. One of these cases was between orange and yellow, and the other case was between orange and red. While this result may not directly imply that drivers would have difficulty discriminating between traffic control signs representing these combinations of colors and sheeting types, the result may indicate potential for color confusion at these two color boundaries.

It is likely that the participants in the experiment could have identified differences between these two pairs of sign colors and materials if the participants had viewed the pairs simultaneously side-by-side. However, recognition of traffic control signs should not require simultaneous presentation but should be unambiguous at a distance at which drivers may acquire textual information. In this sense, the method of direct perceptual hue and saturation rating employed in the experiment is closer to the absolute identification of the colors of individual signs required of most driving tasks. The results of the experiment suggest that caution should be applied in any future deliberations to alter the shapes and sizes of the red, orange, and yellow color areas specified for traffic control signs. The results also indicate that there is no pressing need to modify the FHWA color areas at the present time, but improvements may be possible so as to enhance separation of those color boundaries that are more difficult to discriminate.

As concerns perceptual measures of brightness, the diffuse white reflector with a color filter was always judged the brightest for any given color. Retroreflective properties tended to produce darker colors relative to this reference condition. Instrument measurements of luminance confirmed this trend. This overall reduction in luminance that accompanies retroreflective properties may partially explain the loss of saturation of these colors when compared to the diffuse reference condition.

The experiment employed the basic method of direct color scaling developed by Abramov et al.(5) This method does not require precise equipment and controlled viewing conditions and lends itself readily to field applications. The present experiment represents a validation of the basic technique using actual samples of highway traffic sign materials under variable field viewing conditions (differing degrees of cloud cover, passing clouds, differing sun angles, etc.). The technique performed well under such field conditions, producing color areas that were consistent and orderly in the UAD color space.

Gordon et al. expanded on this technique, and the present experiment incorporated some of the enhancements.(6) However, although they noted that about 5 percent of the research participants did not use the percentage scales appropriately, Gordon et al. claimed that research participants did not need special training to use the method. To the contrary, the present researchers found that substantial training was necessary beyond simply administering the formal verbal instructions (see appendix A). A partial explanation of this discrepancy may lie in the differences between the two samples of research participants. Most of the participants in the original studies were somewhat experienced, having participated in several experiments, and rather homogenous, coming from an academic environment.[6] The present experiment employed a broader sample of naïve research participants recruited from the general driving public.

With the caveat that additional training may be necessary, the present experiment may be regarded as a validation of the direct color scaling technique applied to a broader sample of research participants and testing conditions. The results of the experiment indicate that the basic color scaling technique can be employed successfully in an outdoor environment to provide technical answers to practical problems in highway and traffic engineering. The problems investigated in the experiment revolve around determining the size and shape of the FHWA color areas used to specify traffic control signs. Future implementations of this technique might be used to answer specific engineering questions about particular combinations of color and retroreflective properties. The experiment also added an analogous brightness scaling technique, which proved equally successful. The researchers recommend that these methods be considered in future research to address questions concerning driver perception of the color properties of traffic control signs.

APPENDIX A. PARTICIPANT MATERIALS

This appendix contains instructions and materials provided to study participants during training and data collection. In some cases, representative samples have been provided instead of the full set of materials.

INSTRUCTIONS FOR SCALING HUE, SATURATION, AND BRIGHTNESS

As each color sample panel is mounted in the tripod, describe your sensation of the light reflected from this particular color sample. First, divide your sensation into three parts. The first part consists of hue. You must divide your sensation of hue into red, yellow, green, or blue. These are the only words you may use to describe the hue. If you wish, you may use pairs of names to describe a hue. When you are describing your sensations of hue, please state them in percentages. If there is any hue at all, the percentages you assign to these words must add up to 100 percent. For instance, your sensation might be 40 percent red and 60 percent yellow or 86 percent green and 14 percent yellow or 95 percent blue and 5 percent red, and so on. It can also be 100 percent of one of the hues. The only limitation is that you may not pair red with green, nor pair yellow with blue; all other combinations are allowed. If there is no hue at all and the color is either black or white or some shade of gray between black and white, then the percentages you assign to the words will be 0 percent and they will add up to 0 percent.

Think carefully about your answer and try to be as precise as possible. Remember that the term *hue* refers to your own sensation elicited by the light reflected off the material. You are not being asked how you might create the particular hue you saw. You are being asked to describe your sensation. Please note that there are no right or wrong answers; you are simply describing your sensation.

The second part of your sensation is not hue but is related to hue. After describing the hue of your sensation, you must consider what percentage it formed of your entire sensation; that is, what is the percentage of chromatic versus achromatic-plus-chromatic sensations? This value is called *apparent saturation*. It refers to the strength or concentration of hue in your total sensation. A total absence of hue would be represented by 0 percent saturation, in which case a color would be white or gray or black. At the other extreme, a fully saturated color with 100 percent saturation would be as far away as possible from white or gray or black. Think of your total sensation when you see a color sample as something contained in a bucket. Now pour a little bit of a hue into the bucket and stir. What has happened to the saturation? Now add a little bit more and stir. Again, how has saturation changed?

The third part of your sensation is neither hue nor saturation. This part is achromatic; it is not sensitive to color. In this case we want you to rate the brightness of the light reflected off the material. Brightness is the sensation of the amount of light. Brightness is expressed on a scale from bright to dark, with 100 percent being the brightness of the sky on a clear day, and 0 percent being the brightness of a completely dark surface that does not appear to reflect any light at all. Zero percent brightness is what you would perceive in a completely dark room with no light whatsoever.

First, you will scale some color patches printed in a training booklet in order to get used to the task. Next, you will scale a few actual color sample panels as practice trials. This training is designed to familiarize you with the procedure. We will tell you when they end and the experimental trials begin. Do not leave any blanks on your answer sheet. It is always better to guess. The first data collection period after the practice trials will last about 30 min. Then you will have a break. Do you have any questions?

INSTRUCTIONS FOR BRIGHTNESS RANKINGS

In this part of the experiment you will rank order the brightness of five color sample panels lined up in a row with the letters A–E under each panel. All of the panels will be of a similar color. On your answer sheet, please rank the five panels for their sensation of brightness only. Place the letter of the brightest sample under the number 1 on the answer sheet, and the letter of the next brightest sample under the number 2, and so on, until you have ranked all five samples. When you are done, the letter corresponding to the brightest sample will be under the number 1, the letter corresponding to the darkest sample will be under the number 5, and the other letters will be arranged in between according to their perceived brightness.

SATURATION SUPPLEMENT

If this is still confusing, consider pink and red. Red is a highly saturated color; it would be given a high saturation score. However, if red becomes more white or gray, it becomes pink and would be given a lower saturation score. With even less saturation, it becomes a pastel pink. If the color continued changing in this manner, at some point it would turn white or gray and would get a saturation score of 0 percent. Remember that black would also have a saturation score of 0 percent.

TRAINING EXAMPLES AND PARTICIPANT PRACTICE

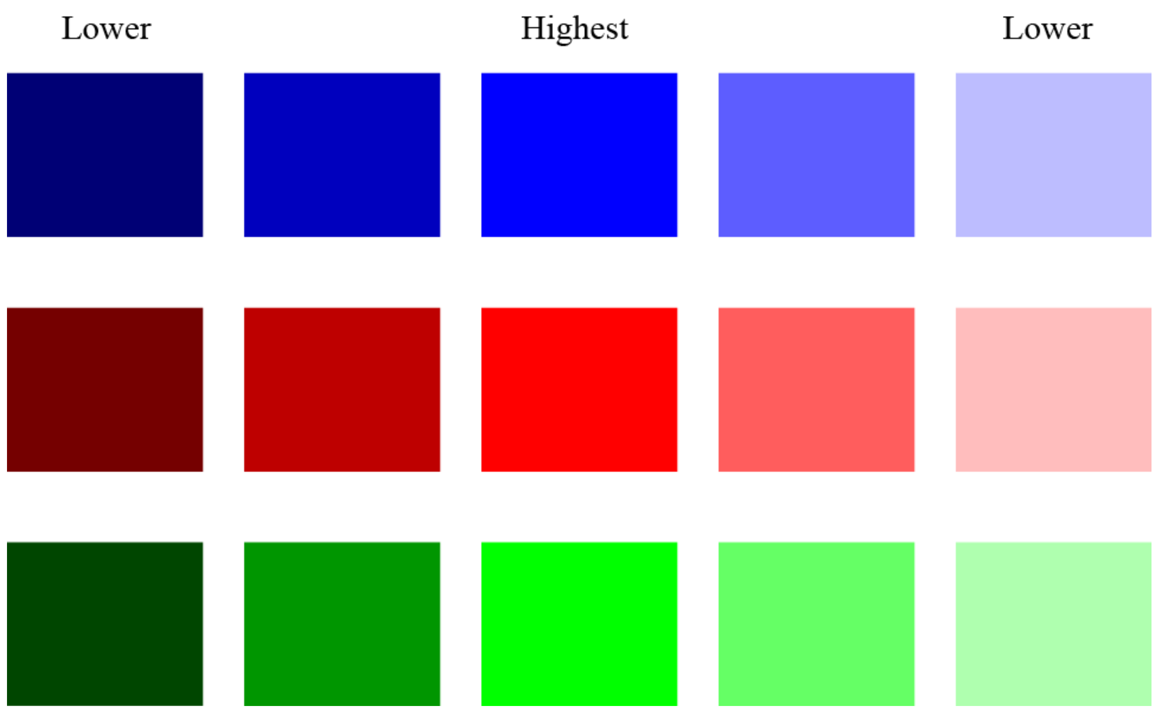


Figure 24. Illustration. Training examples for saturation.

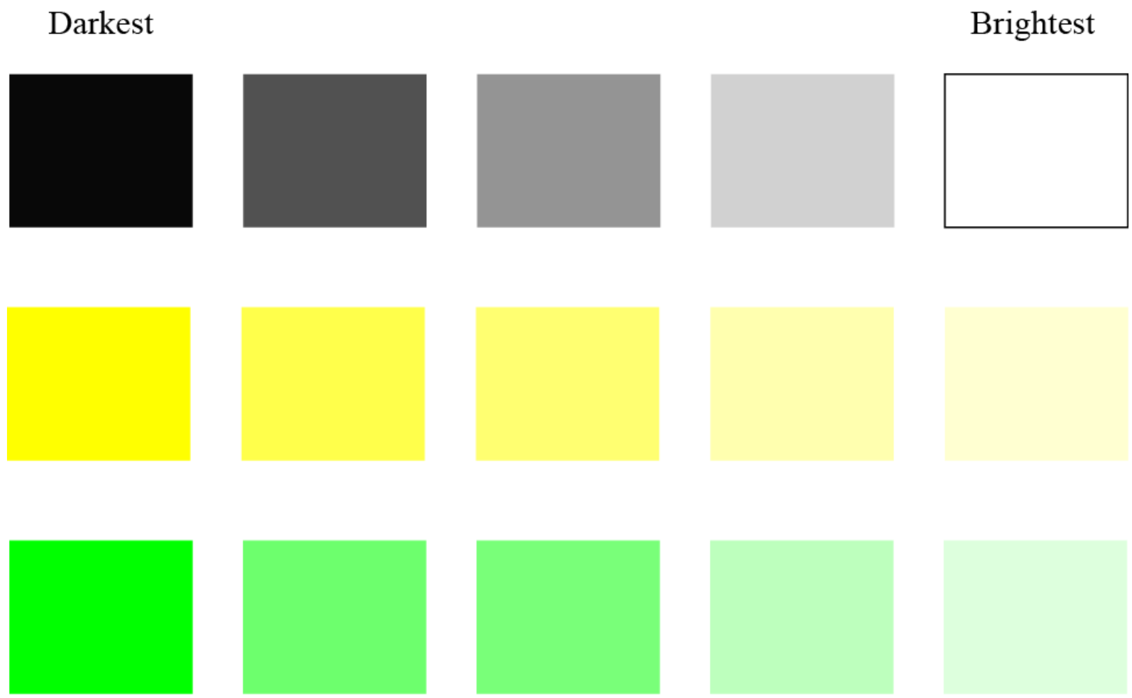


Figure 25. Illustration. Training examples for brightness.

Twenty-four color samples were provided to participants for practice in assigning percentages for the red, green, yellow, and blue in the sample as well as for saturation and brightness. An example from the participant practice is provided in figure 26.

Red	Green	Yellow	Blue	Saturation	Brightness
______%	______%	______%	______%	______%	______%

Figure 26. Illustration. Example participant practice sample.

EXPLANATION OF THREE COLOR SCALES

Hue = What color or colors are present in the sample, even in small amounts?

Saturation = How concentrated or strong is that color or color combination?

Brightness = If all color were to disappear, how bright or dark would the sample be?

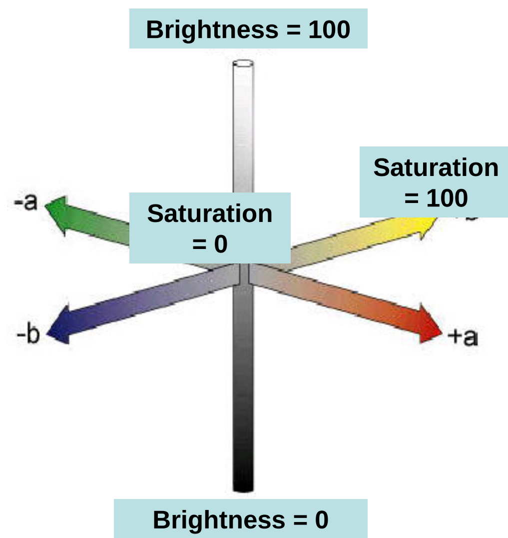


Figure 27. Illustration. Color dimensions.

PARTICIPANT RESPONSE SHEETS

The response sheets provided to participants were in a tabular format with space to enter data for red, green, yellow, blue, saturation, and brightness percentages for each sample. Space was also provided for numbers assigned during the brightness ranking task. Figure 28 shows a representative portion of the response sheets.

Block	Trial						
		RED	GREEN	YELLOW	BLUE	SATURATION	BRIGHTNESS
A	**1**						
	2						
	3						
	4						
	5						
	6						
	7						
	8						
	9						
	10						
		Brightness Ranking		**1= Brightest**	**5= Dimmest**		
		1	2	3	4	5	
		Sequence					

Figure 28. Illustration. Response sheet sample.

APPENDIX B. PHYSICAL MEASUREMENTS

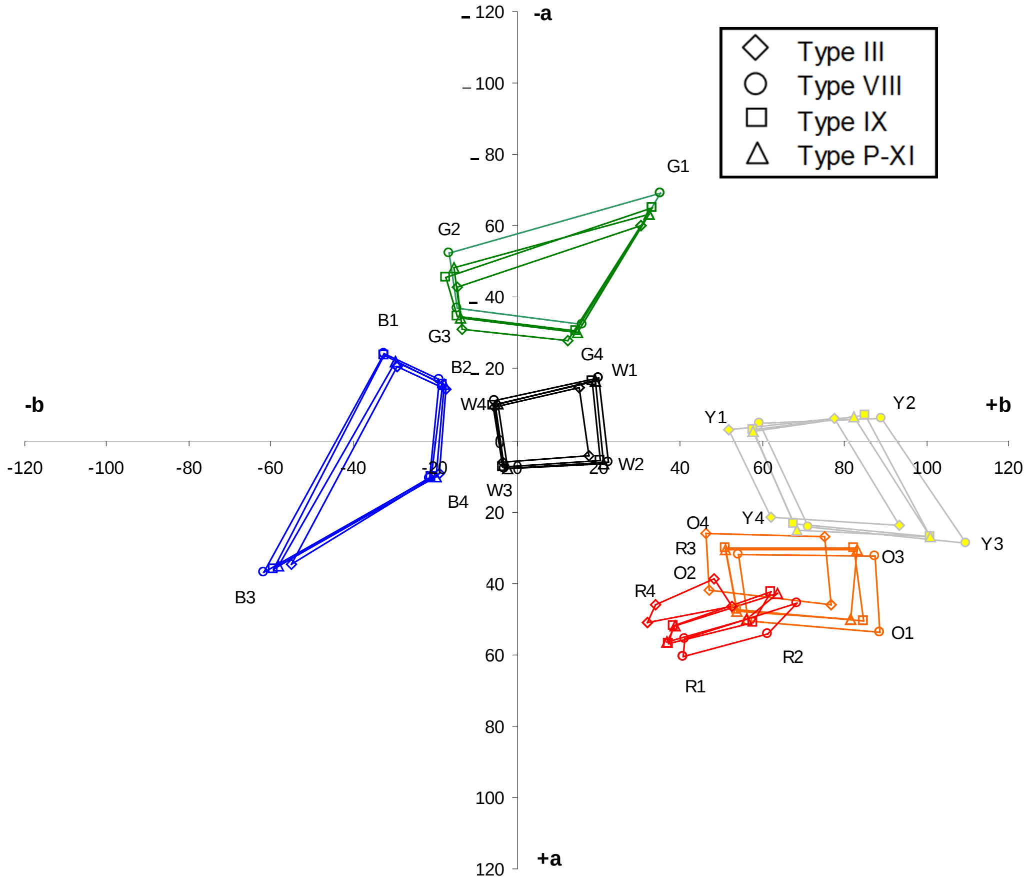


Figure 29. Graph. Laboratory (LabScan® XE) physical color measurements of four retroreflective sheeting types.

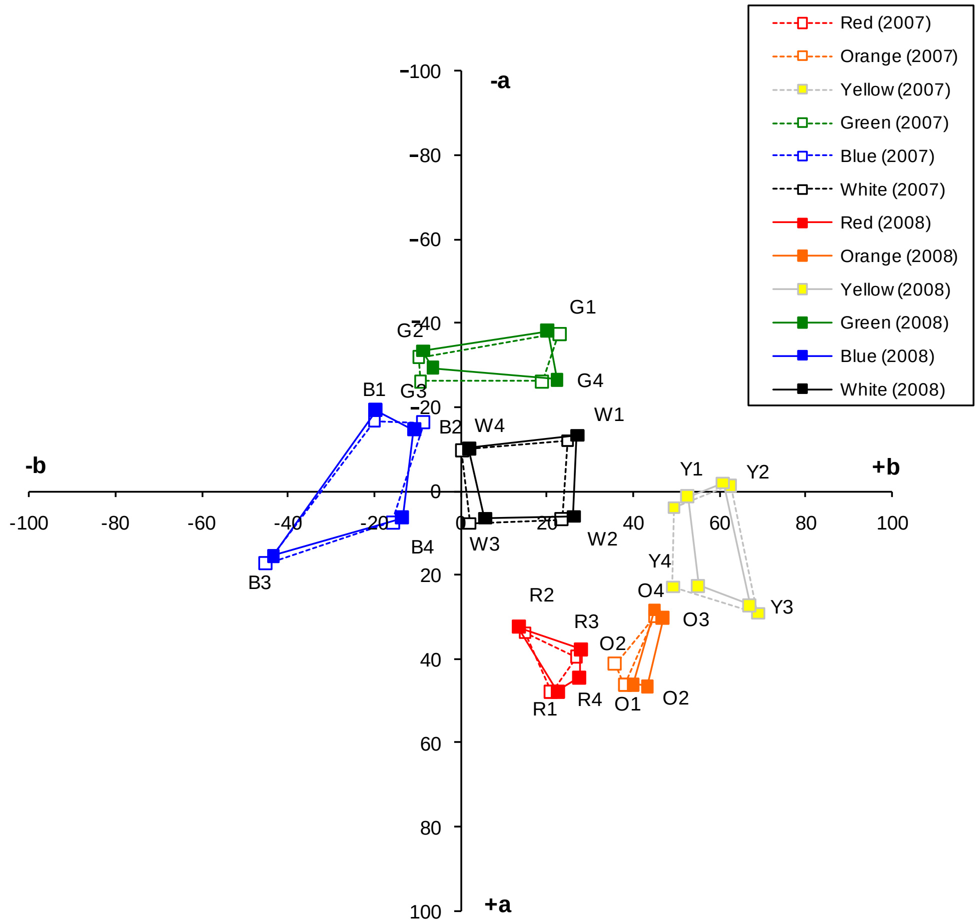


Figure 30. Graph. Comparison of mean field (PR-650) physical color measurements from 2007 and 2008 averaged over four retroreflective sheeting types.

APPENDIX C. HUMAN PSYCHOPHYSICAL RESULTS

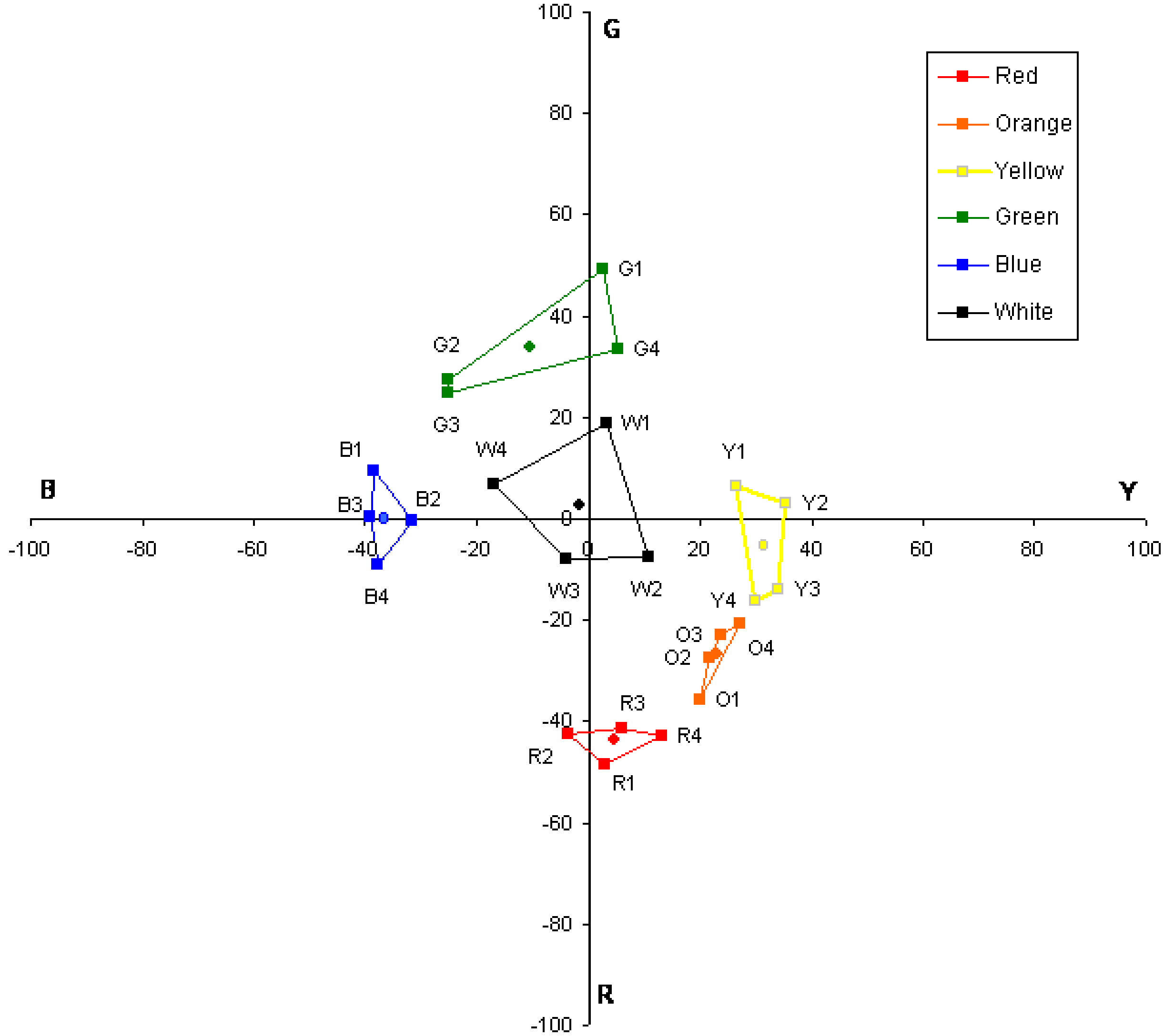


Figure 31. Graph. Mean perceptual color ratings for type VIII sheeting for 17 participants.

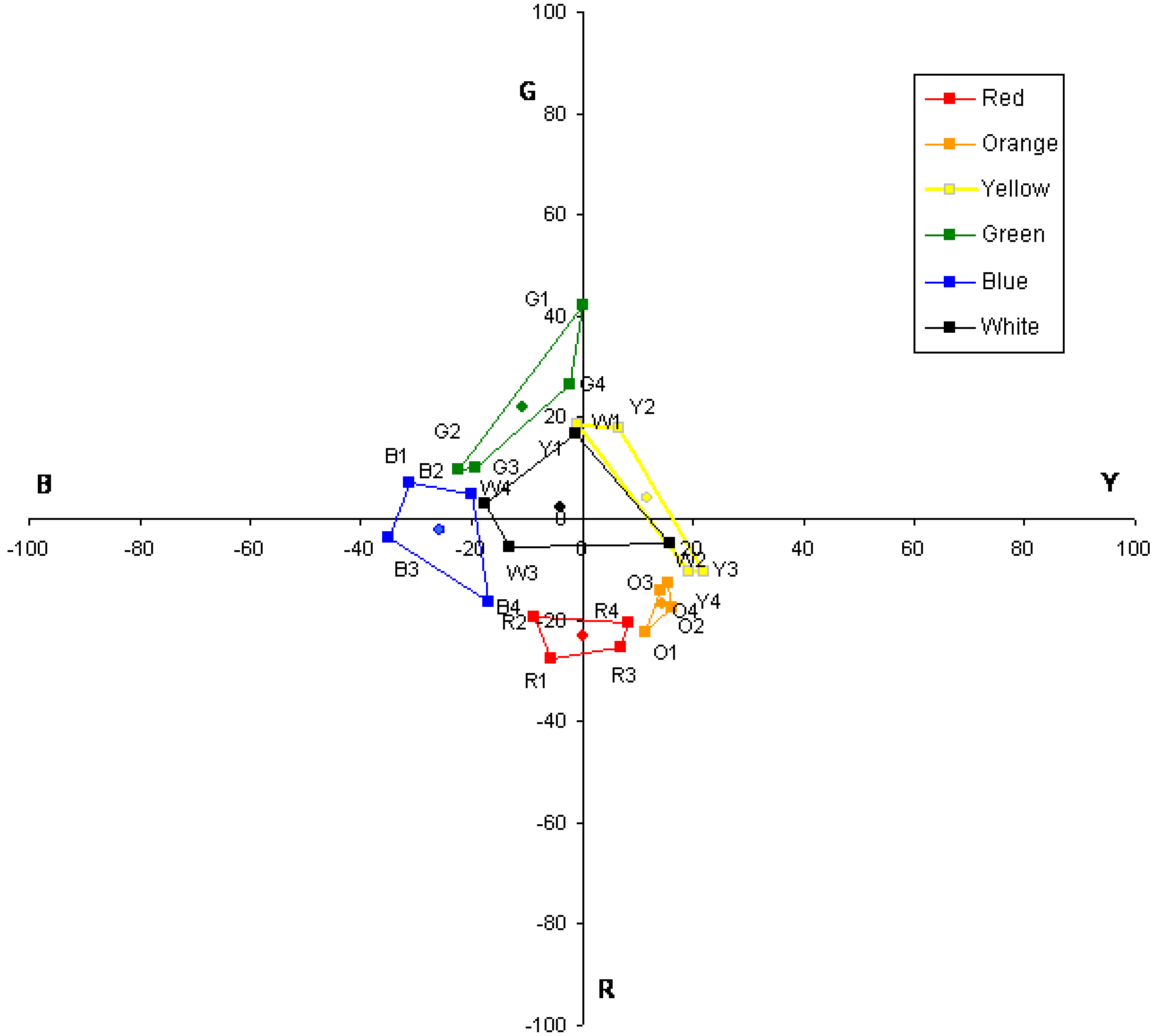


Figure 32. Graph. Mean color ratings for type VIII sheeting for participant 1.

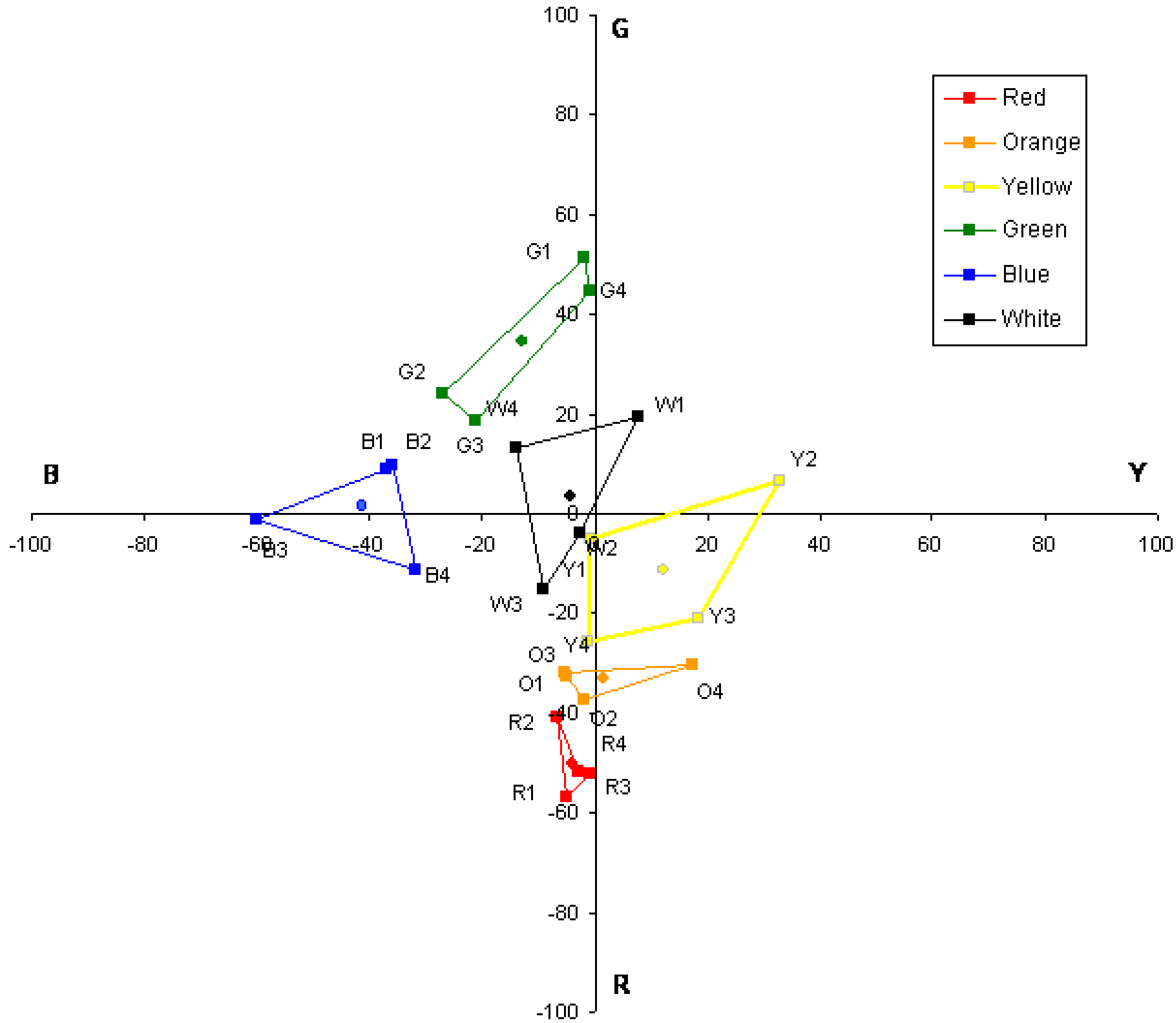


Figure 33. Graph. Mean color ratings for type VIII sheeting for participant 2.

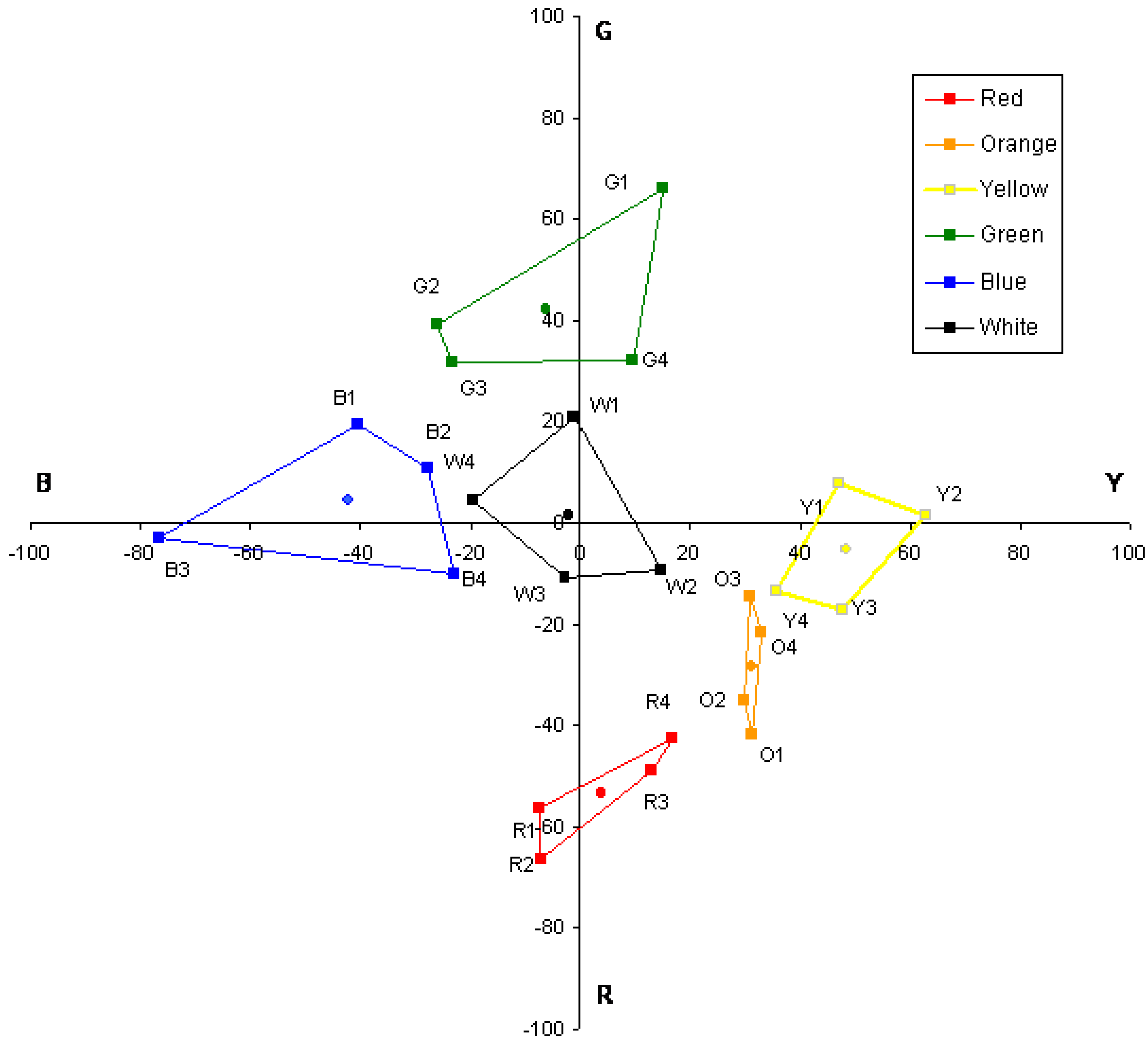


Figure 34. Graph. Mean color ratings for type VIII sheeting for participant 3.

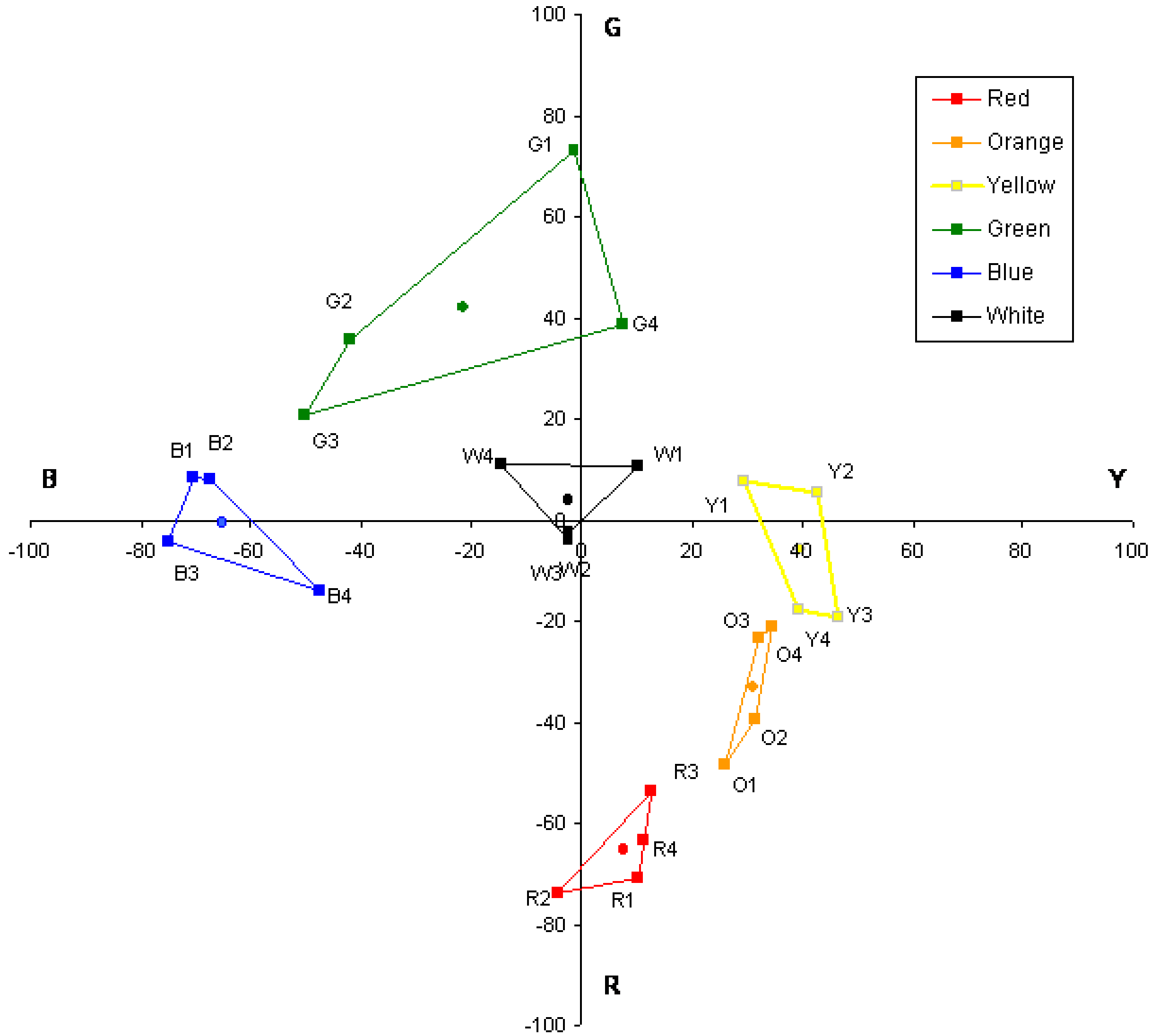


Figure 35. Graph. Mean color ratings for type VIII sheeting for participant 4.

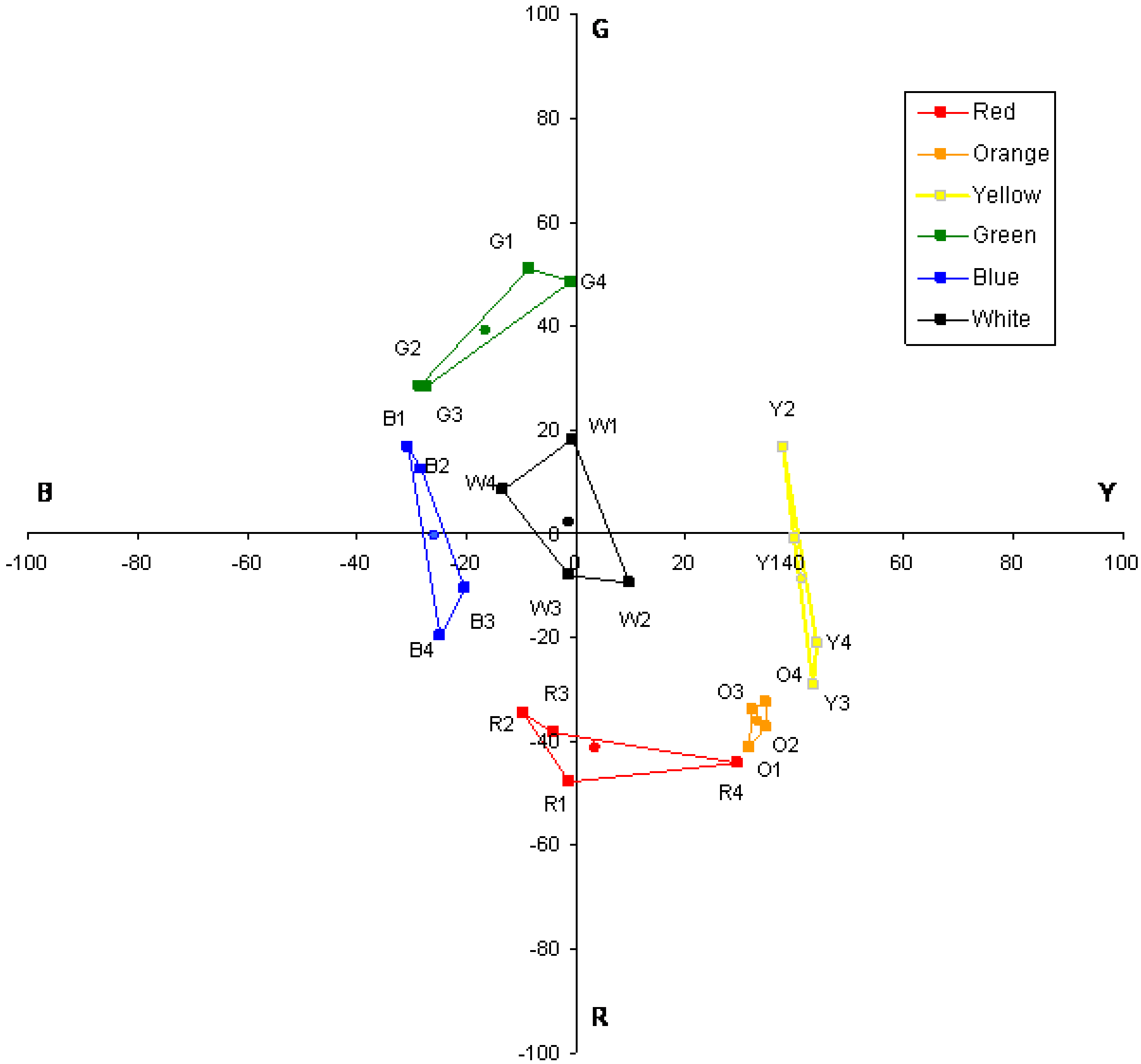


Figure 36. Graph. Mean color ratings for type VIII sheeting for participant 5.

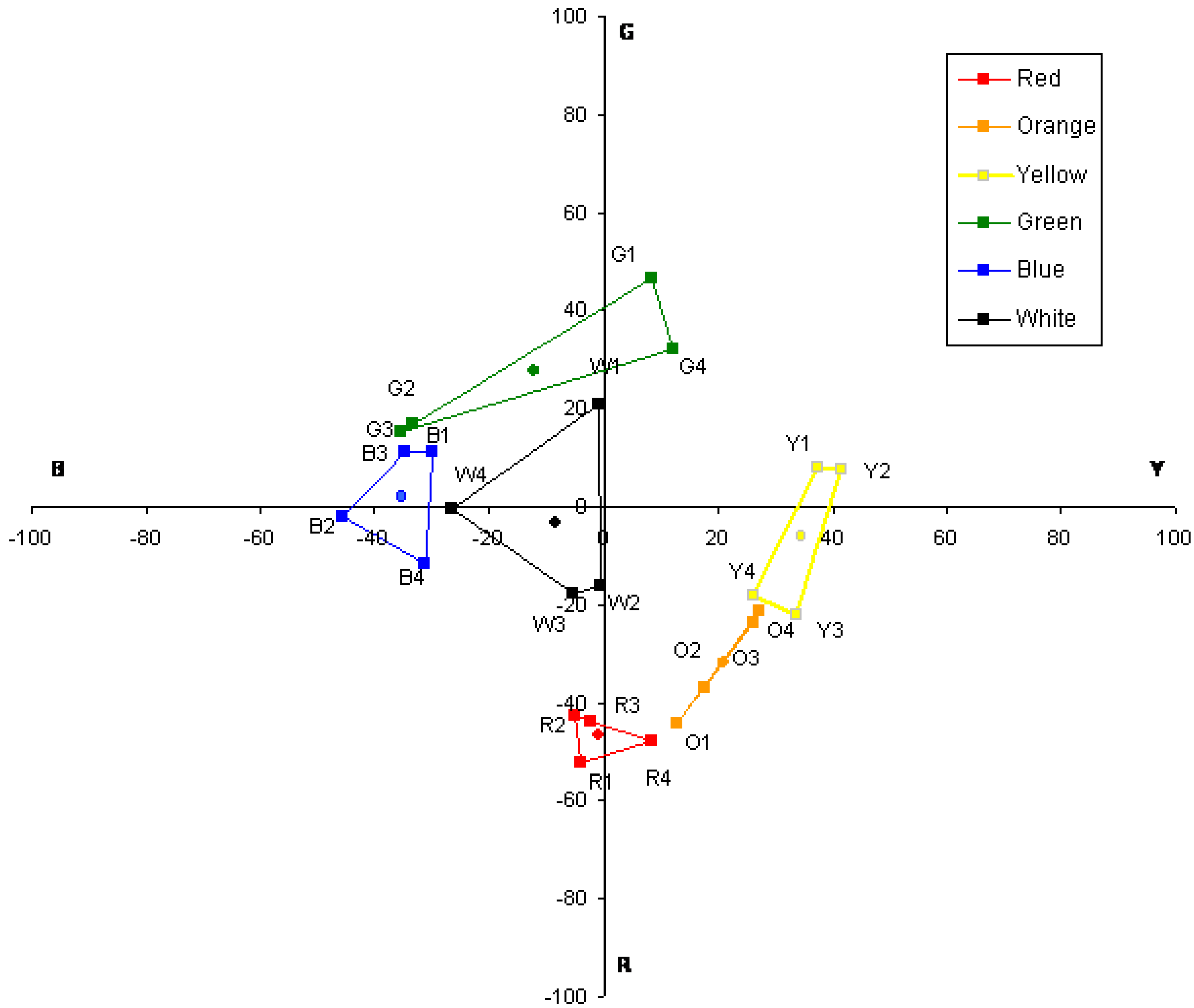


Figure 37. Graph. Mean color ratings for type VIII sheeting for participant 6.

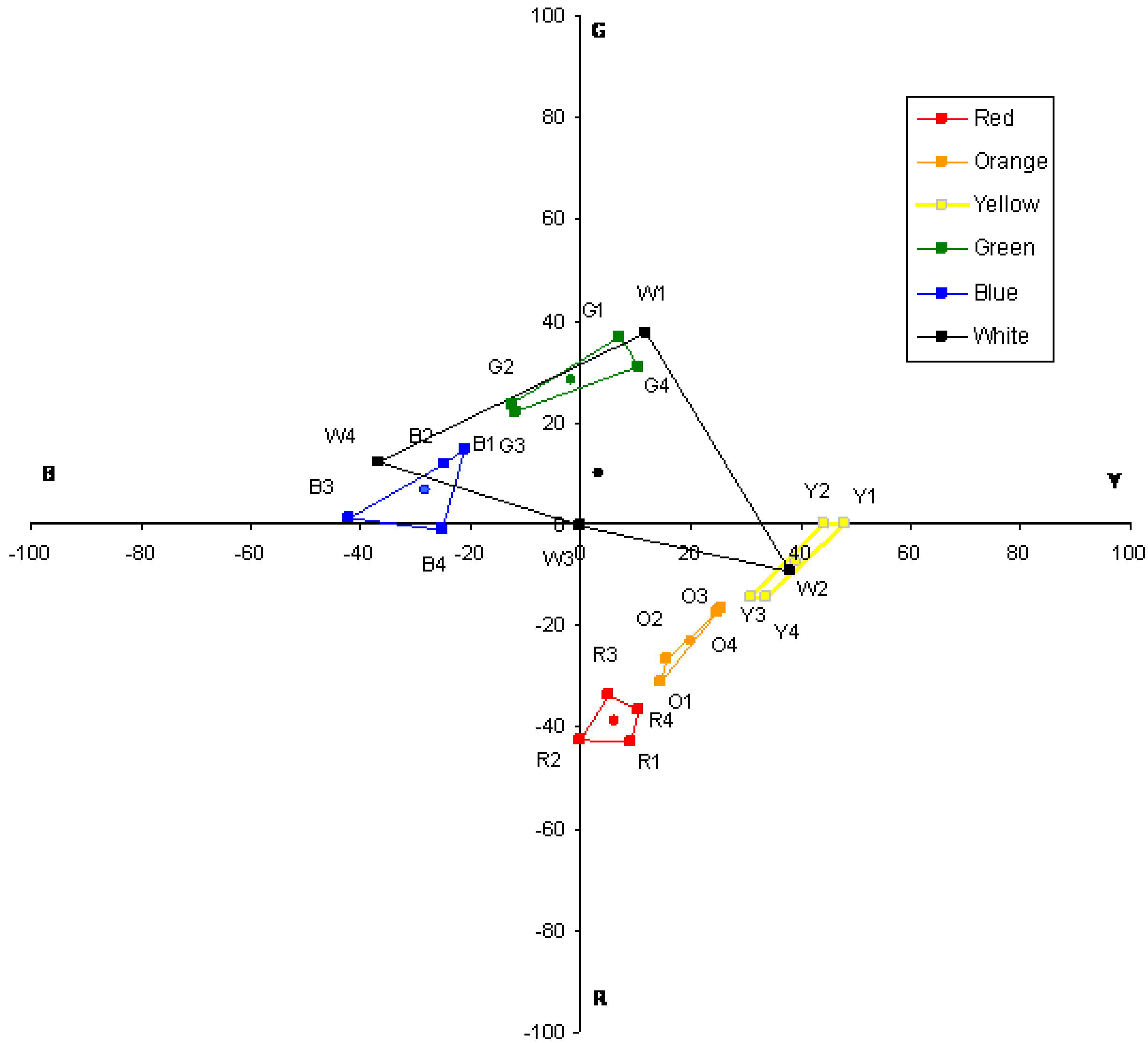


Figure 38. Graph. Mean color ratings for type VIII sheeting for participant 7.

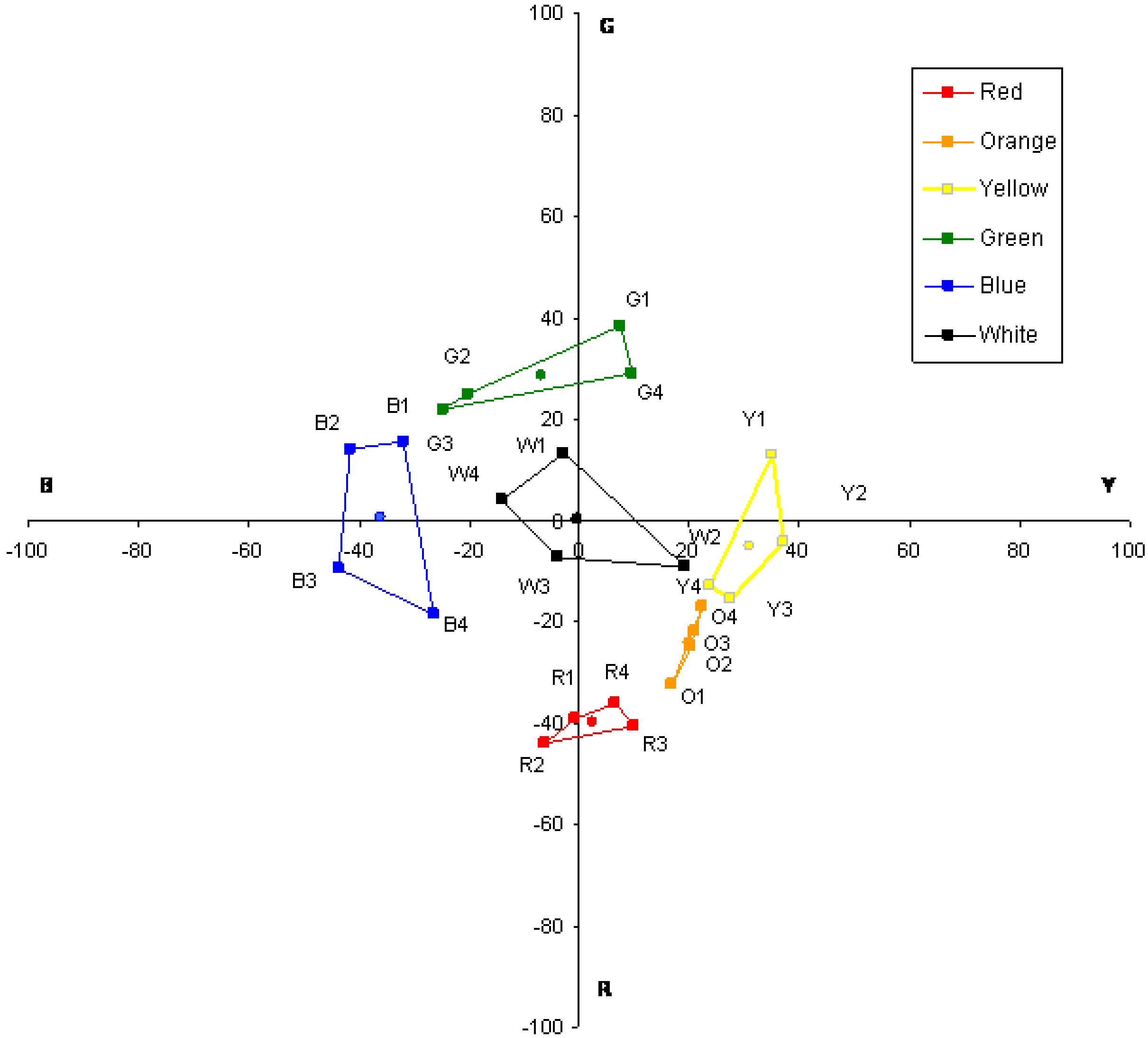


Figure 39. Graph. Mean color ratings for type VIII sheeting for participant 8.

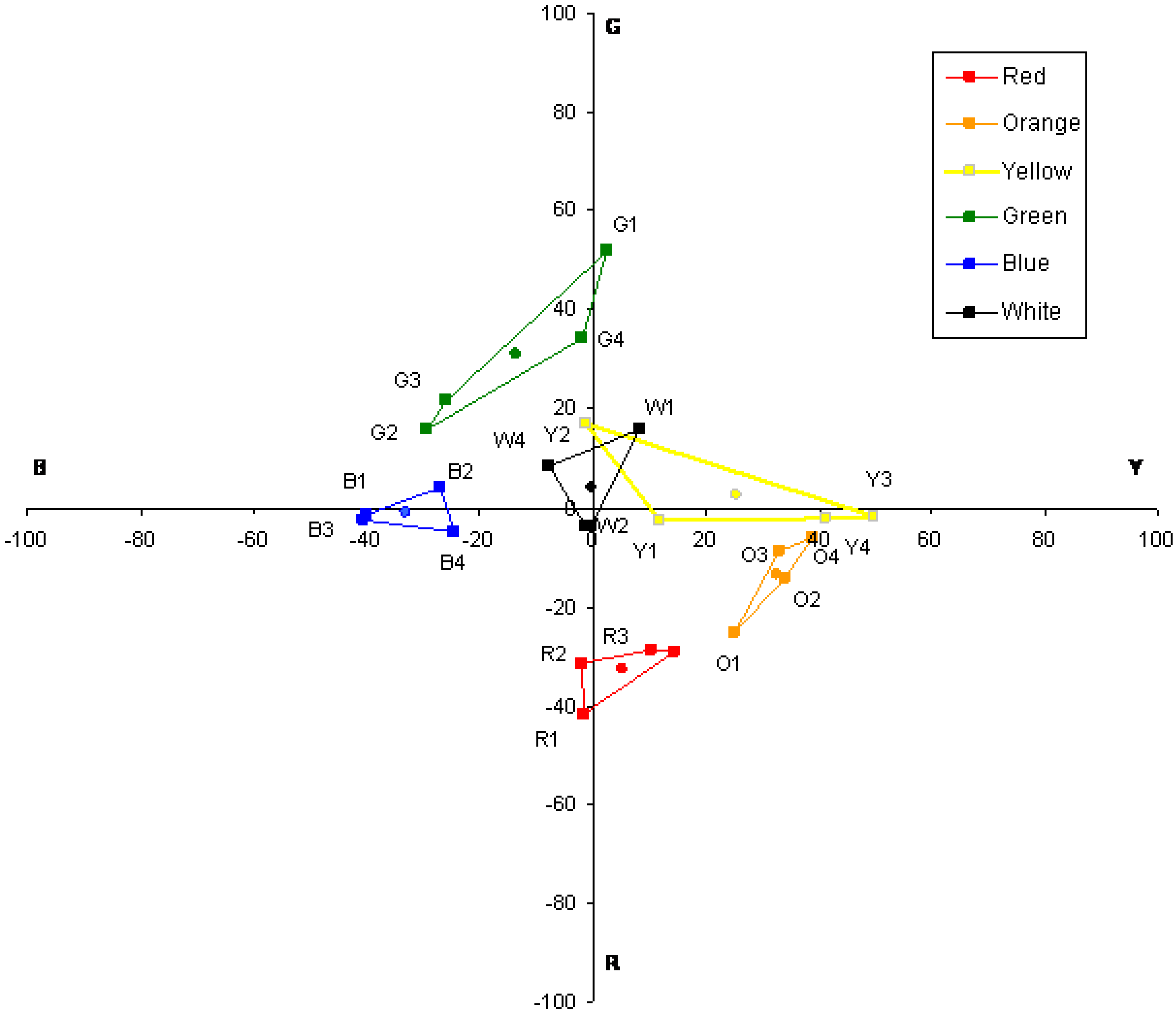


Figure 40. Graph. Mean color ratings for type VIII sheeting for participant 9.

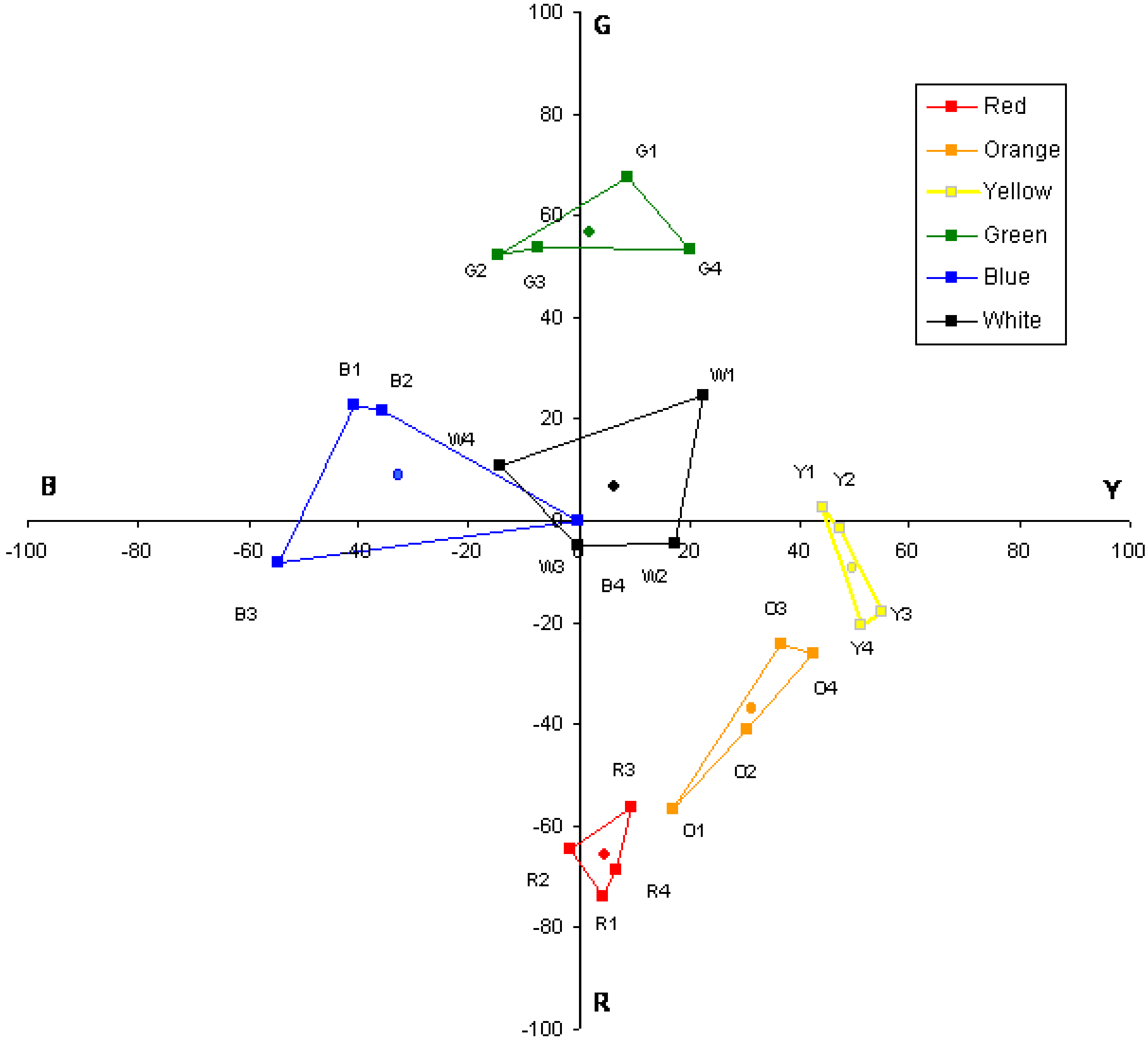


Figure 41. Graph. Mean color ratings for type VIII sheeting for participant 10.

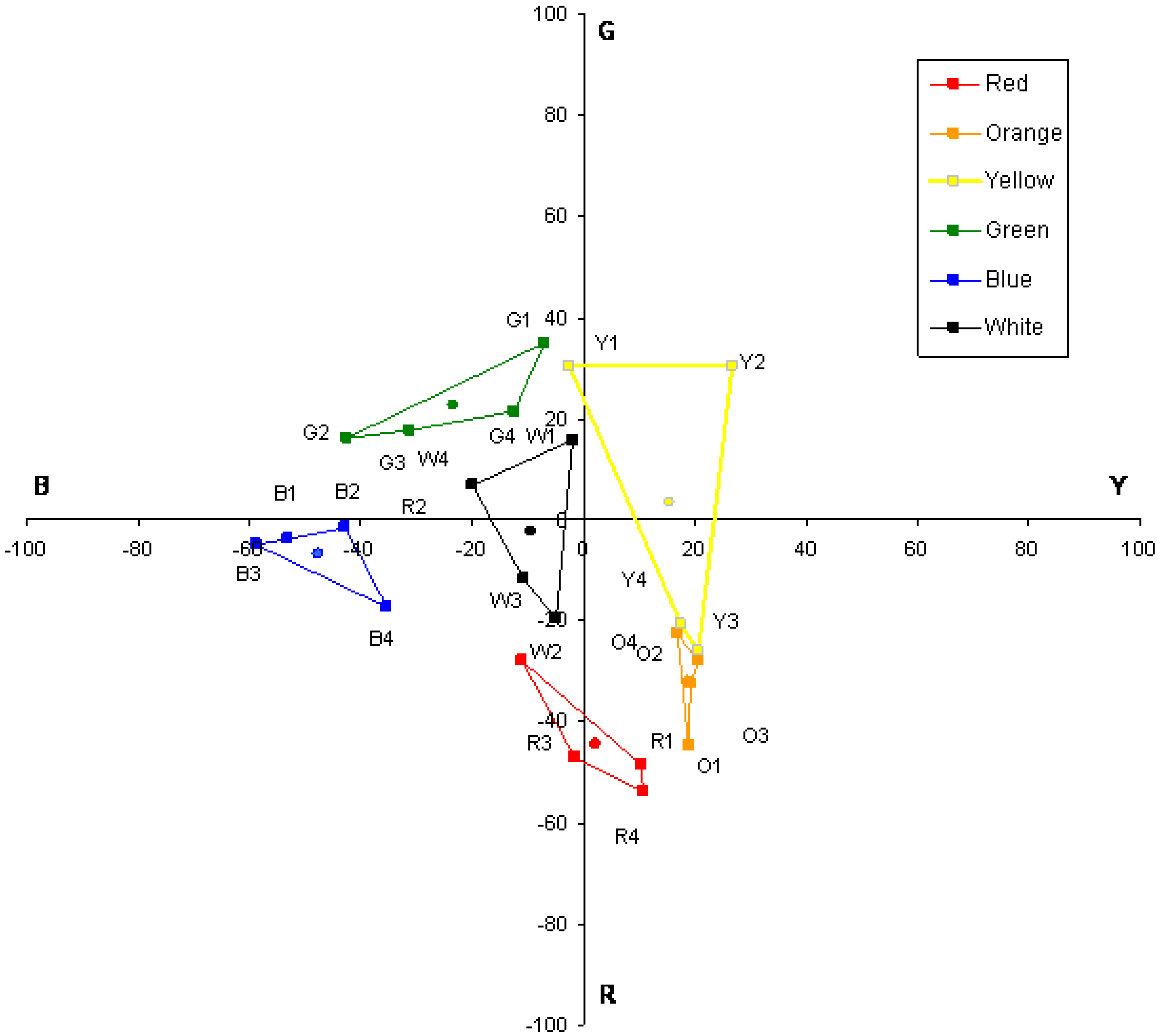


Figure 42. Graph. Mean color ratings for type VIII sheeting for participant 11.

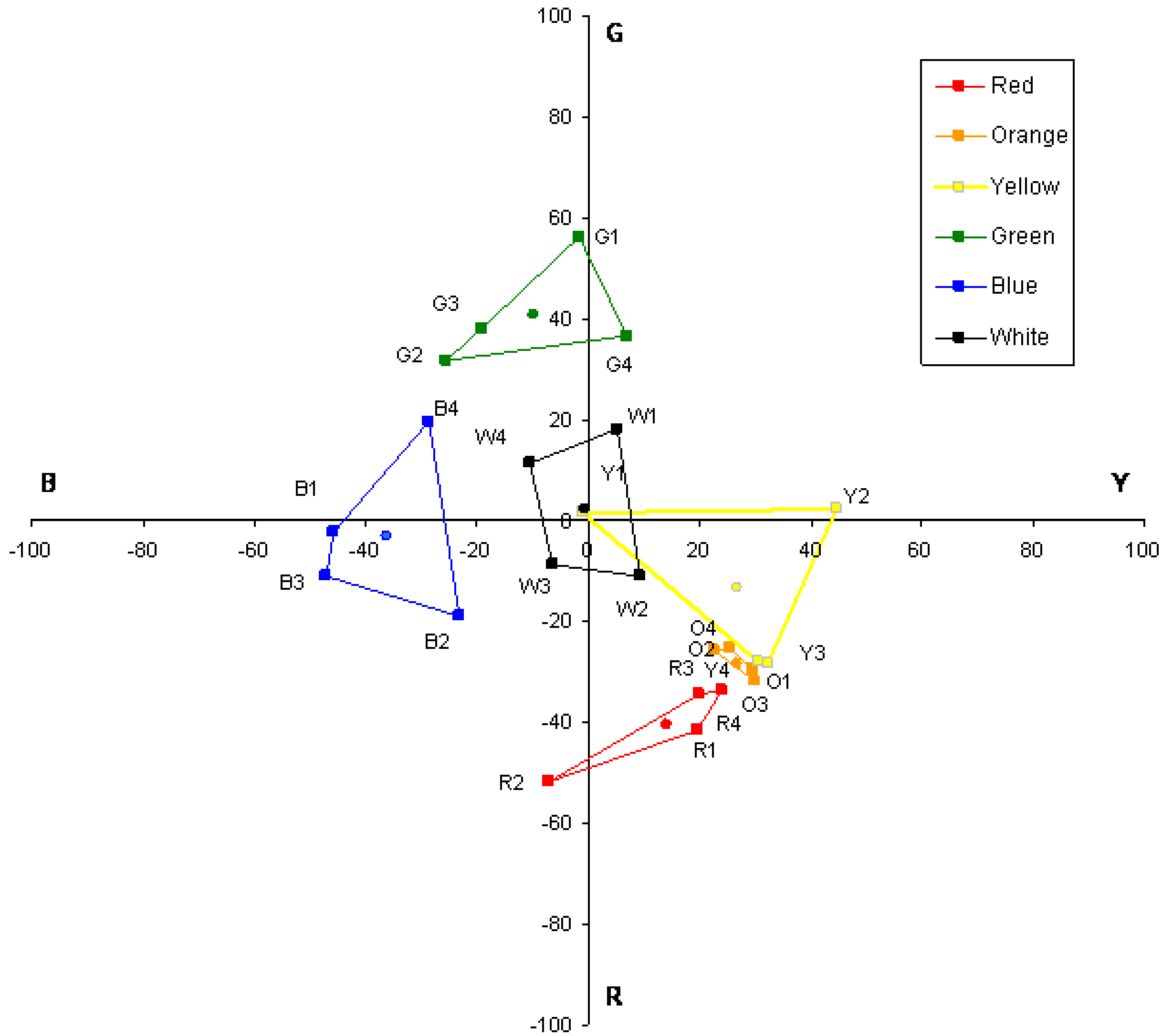


Figure 43. Graph. Mean color ratings for type VIII sheeting for participant 12.

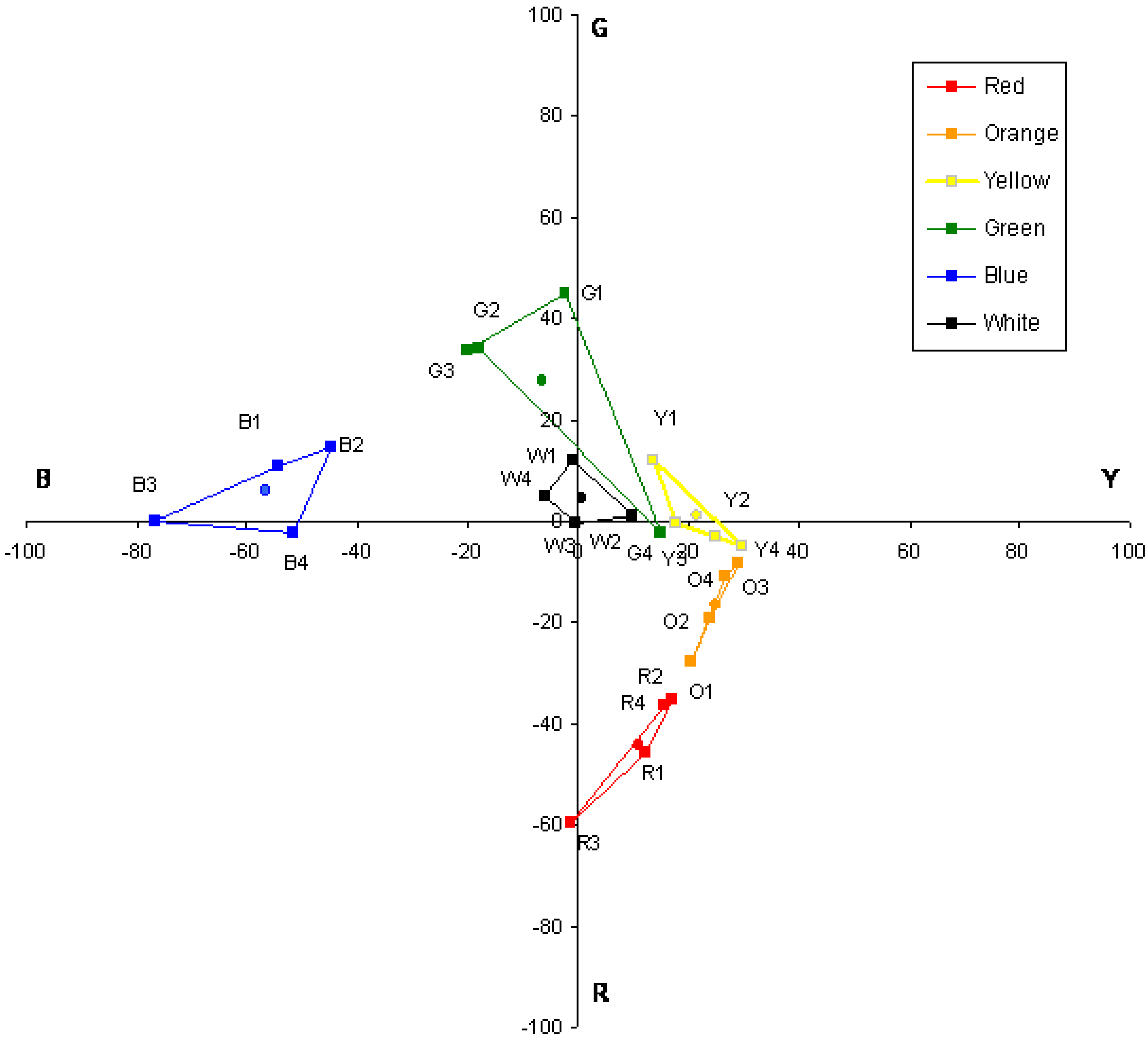


Figure 44. Graph. Mean color ratings for type VIII sheeting for participant 13.

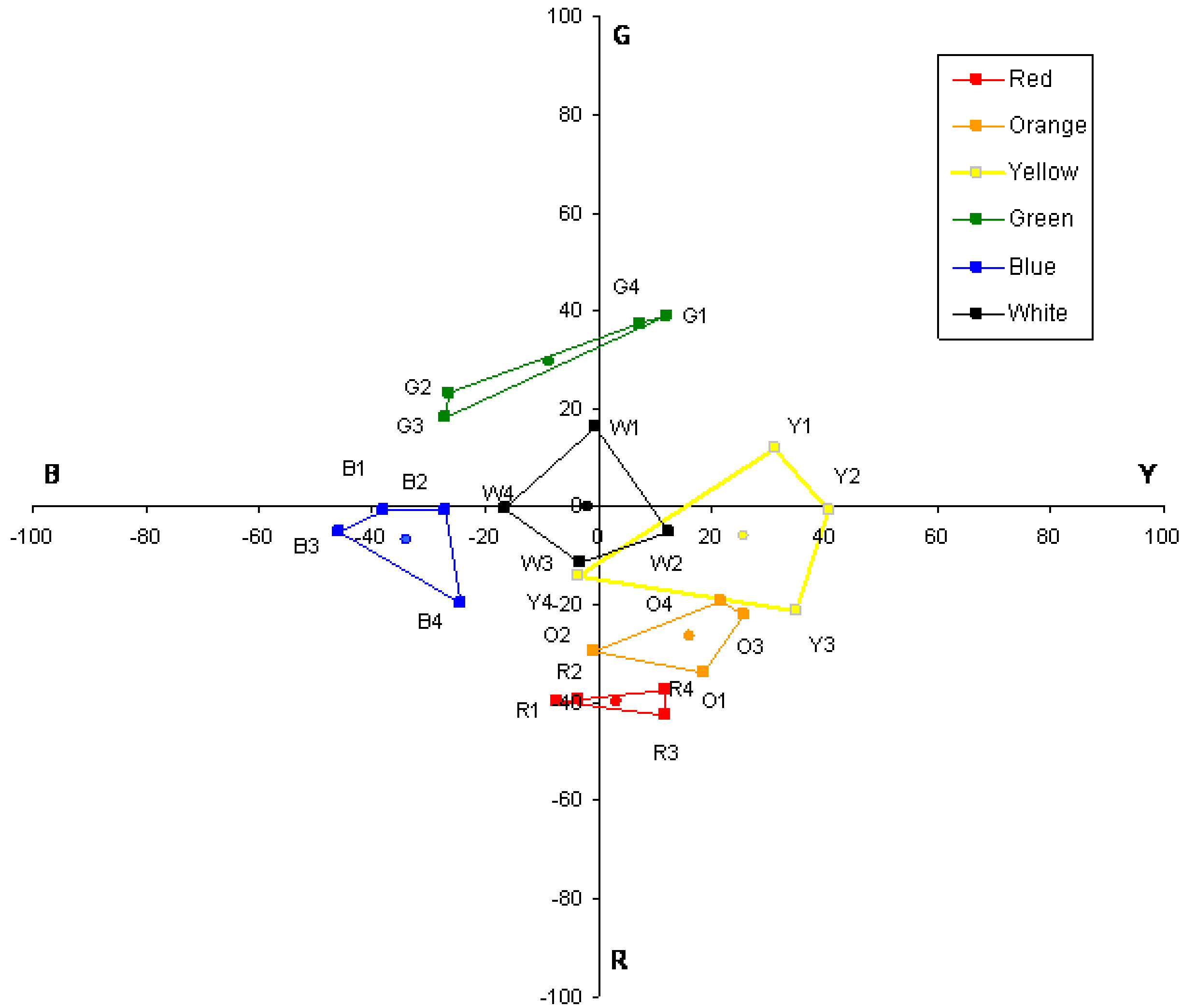


Figure 45. Graph. Mean color ratings for type VIII sheeting for participant 14.

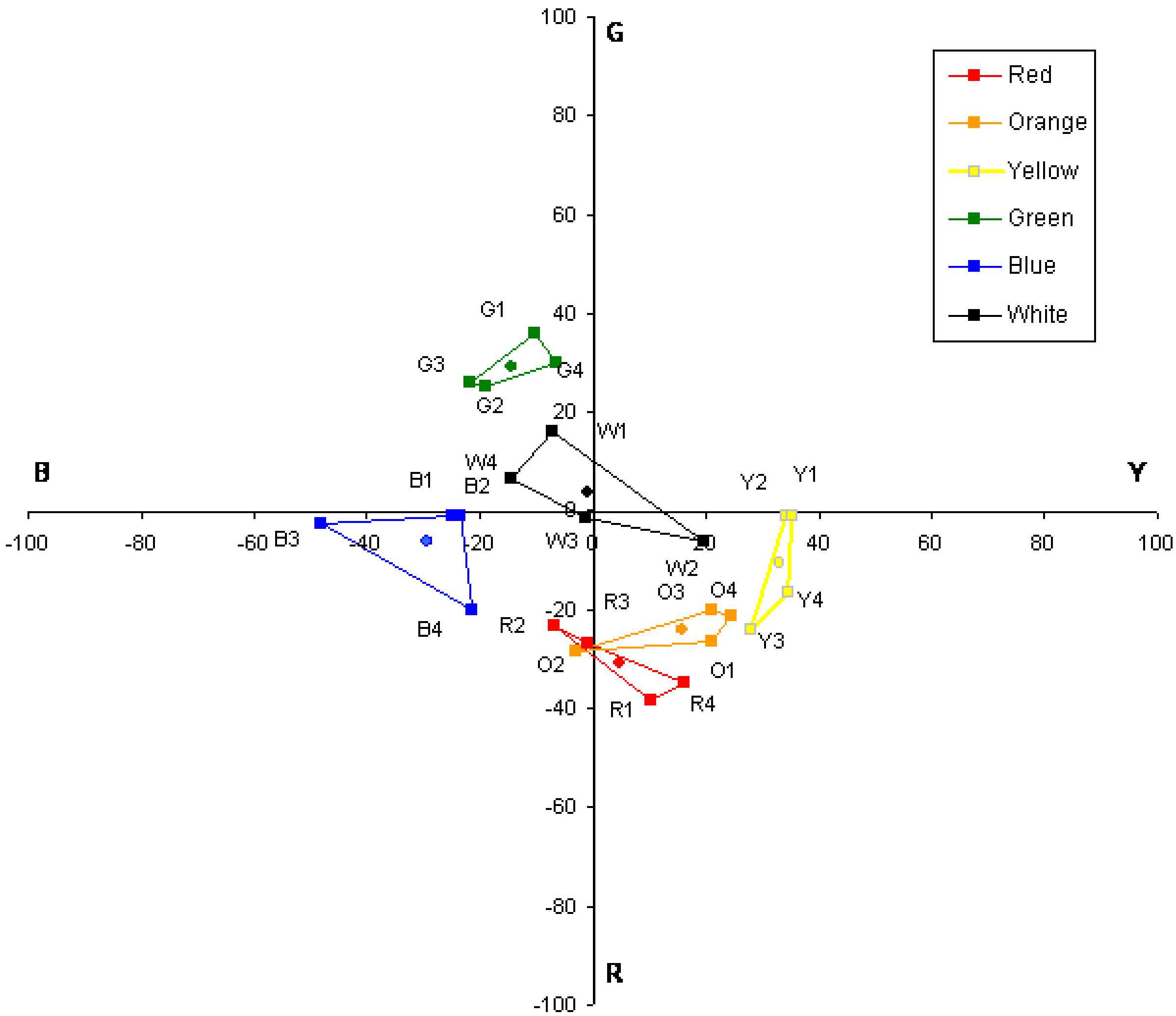


Figure 46. Graph. Mean color ratings for type VIII sheeting for participant 15.

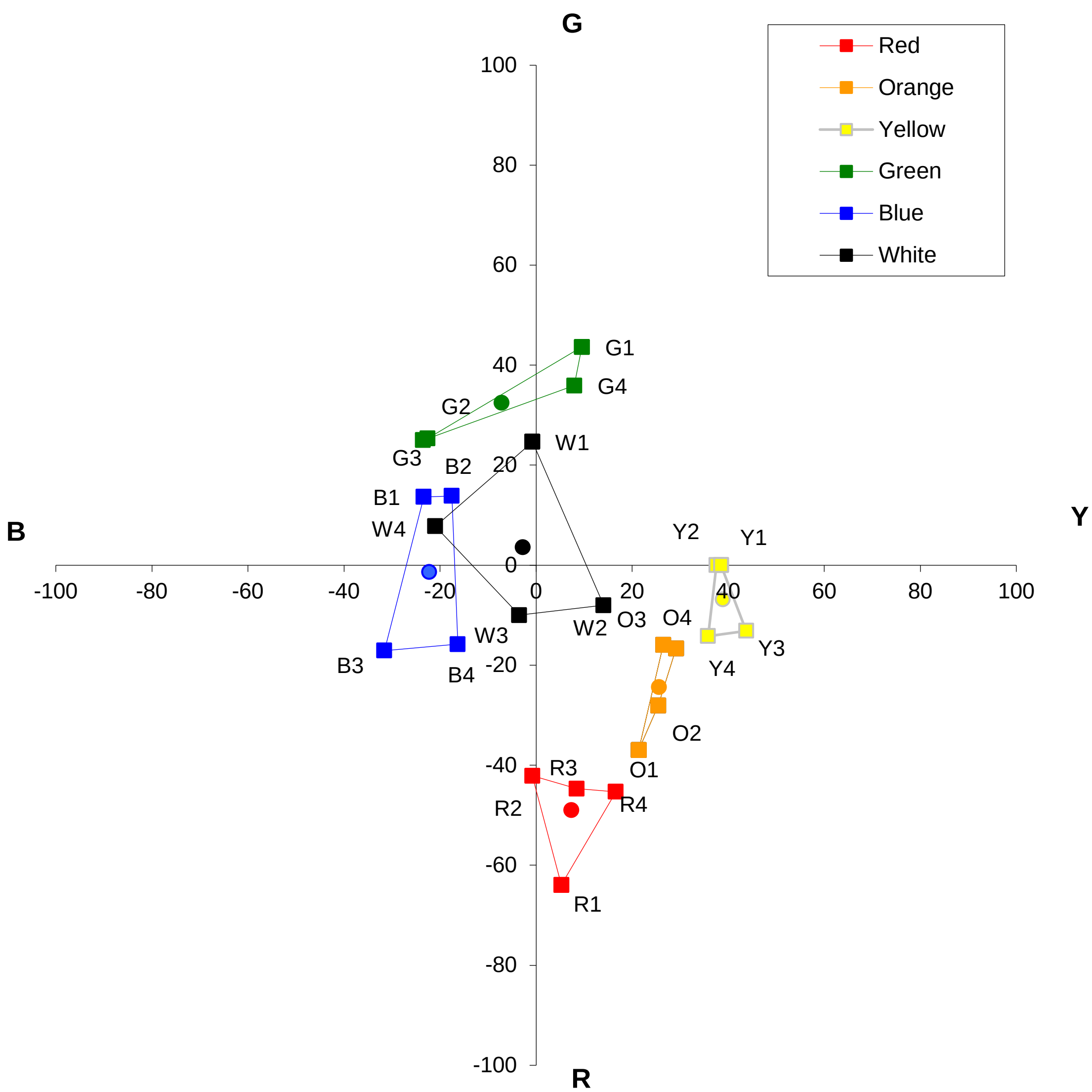


Figure 47. Graph. Mean color ratings for type VIII sheeting for participant 16.

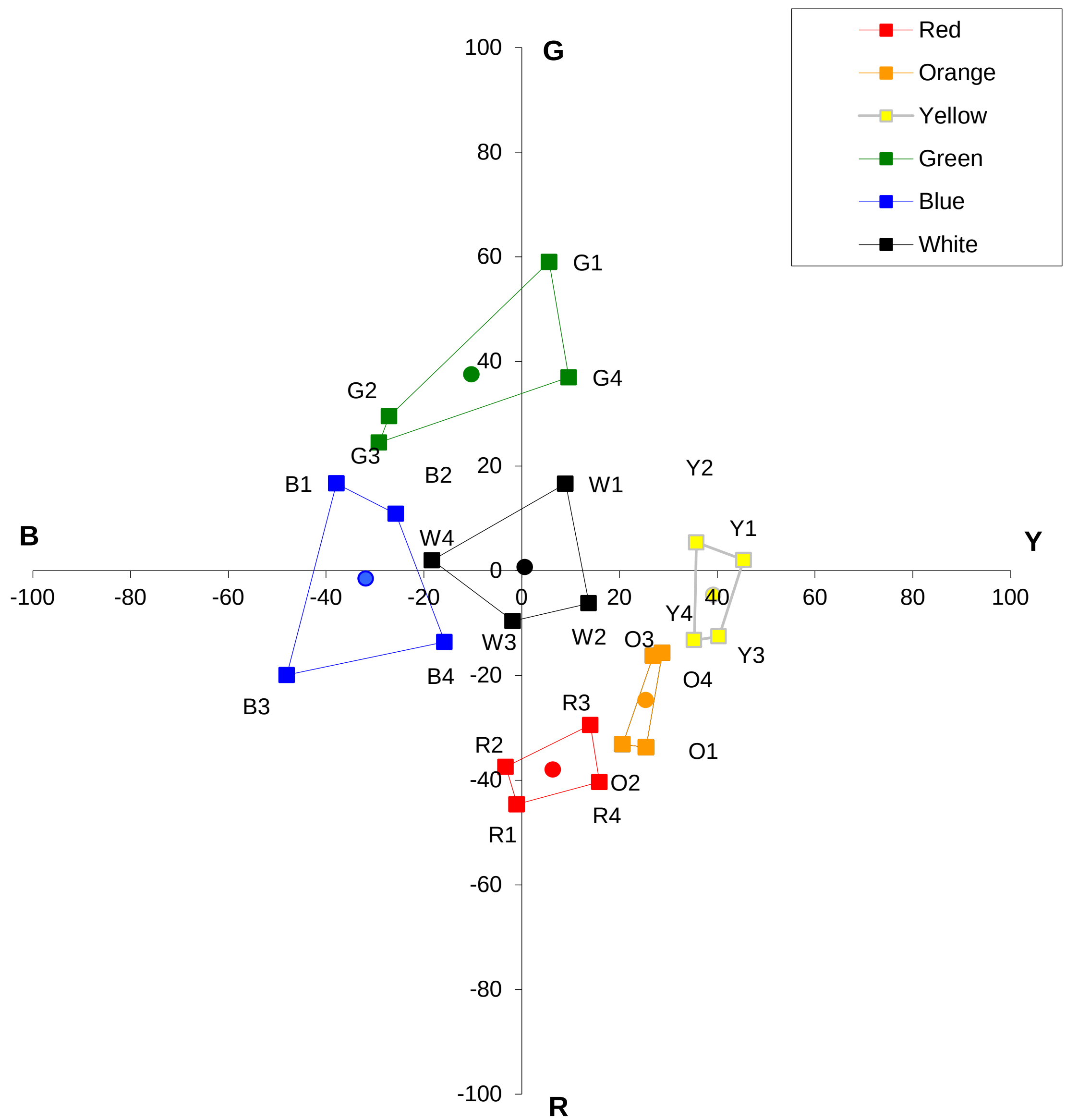


Figure 48. Graph. Mean color ratings for type VIII sheeting for participant 17.

ACKNOWLEDGEMENTS

The FHWA sponsored this work. This research could not have been completed without the assistance of the following personnel: Dana Duke, Lindsey Clark, Megan Hermosillo, and Erin Dagnall. Thomas Granda at FHWA as well as Stephen Fleger and William Perez provided management support for the contract under which the work was conducted.

REFERENCES

1. Johnson, N.L. (1999). “Accuracy in Measuring the Daytime Color of Retroreflective Materials.” *Proceedings of the 24th Session of the CIE*, International Commission on Illumination, Warsaw, Poland.

2. Burns, D.M. (2004). “Daytime Color of Retroreflective Traffic Control Devices: An Investigation into the Relationship between Instrumental Colorimetry in the Laboratory and Field Photometry from the Driver’s Seat.” CIE/USA Annual Meeting.

3. National Archives and Records Administration (2002). “Section 655, Traffic Operations, Appedix to Subpart F.” *Code of Federal Regulations*, 23 CFR § 655, Washington, DC.

4. CIE (1932). *Commission Internationale De l'Eclairage Proceedings, 1931,* Cambridge University Press, United Kingdom.

5. Abramov, I., Gordon, J., and Chan, H. (1990). “Using Hue Scaling to Specify Color Appearance.” *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, pp. 40–51, Santa Clara, CA.

6. Gordon, J., Abramov, I., and Chan, H. (1994). “Describing Color Appearance: Hue and Saturation Scaling.” *Perception & Psychophysics*, Vol. 56, pp. 27–41.

7. Jacobs, G.F. and Johnson N.L. (1995). “Yellow Pavement Markings with Yellow Nighttime Color.” *Transportation Research Record 1495*, pp. 147–155, Transportation Research Board, Washington, DC.

8. Thomas-Meyers, G. and Nagy, A.L. (2003). “Pavement Marking Color Specifications.” *Journal of Vision*, Vol. 3, Association for Research in Vision and Ophthalmology, Rockville, MD.

9. Davis, W., Miller, C., Kennedy, J., and Andersen, C. (2007). “Pilot Study on Daytime Color Appearance of Retroreflective Sheeting.” *Proceedings of the 26th Session of the CIE*, International Commision on Illumination, Beijing, China.

10. CIE (2004). *Colorimetry, 3rd Ed.*, Publication No. CIE 15:2004, Vienna, Austria.

11. ASTM D4956 (2007). *Standard Specification for Retroreflective Sheeting for Traffic Control*, ASTM International, West Conshohocken, PA.

12. ASTM E1349 (2006). *Standard Test Method for Reflectance Factor and Color by Spectrophotometry Using Bidirectional Geometry*, ASTM International, West Conshohocken, PA.

13. Federal Highway Administration (2003). *Manual on Uniform Traffic Control Devices*, Washington, DC.

14. McLellan, M., Lind, L., and Kime, R. (1995). “Hue Angle Determiniations and Statistical Analysis for Multiquadrant Hunter L,a,b Data.” *Journal of Food Quality*, Vol. 18, pp. 235–240.

HRDS-10/04-13(WEB)E