What is white?

J. M. Bosten

Department of Psychology, University of California, San Diego, La Jolla, CA, USA

School of Psychology, University of Sussex, Brighton, UK

R. D. Beer

Department of Psychology, University of California, San Diego, La Jolla, CA, USA

D. I. A. MacLeod

Department of Psychology, University of California, San Diego, La Jolla, CA, USA

To shed light on the perceptual basis of the color white, we measured settings of unique white in a dark surround. We find that settings reliably show more variability in an oblique (blue-yellow) direction in color space than along the cardinal axes of the cone-opponent mechanisms. This is against the idea that white perception arises at the null point of the cone-opponent mechanisms, but one alternative possibility is that it occurs through calibration to the visual environment. We found that the locus of maximum variability in settings lies close to the locus of natural daylights, suggesting that variability may result from uncertainty about the color of the illuminant. We tested this by manipulating uncertainty. First, we altered the extent to which the task was absolute (requiring knowledge of the illumination) or relative. We found no clear effect of this factor on the reduction in sensitivity in the blue-yellow direction. Second, we provided a white surround as a cue to the illumination or left the surround dark. Sensitivity was selectively worse in the blue-yellow direction when the surround was black than when it was white. Our results can be functionally related to the statistics of natural images, where a greater blue-yellow dispersion is characteristic of both reflectances (where anisotropy is weak) and illuminants (where it is very pronounced). Mechanistically, the results could suggest a neural signal responsive to deviations from the blue-yellow locus or an adaptively matched range of contrast response functions for signals that encode different directions in color space.

Introduction

White has special status among colors: It has been called “the mother of all unique hues” (Mollon, 2006). Unique white is experienced as achromatic, or an absence of hue—it is neither reddish, nor yellowish, nor greenish, nor bluish. White is one of the most universal color categories, arising as one of the first distinctions in Berlin and Kay’s (1969) scheme of the evolution of color words in language. In the world color survey by Kay, Berlin, Maffi, Merrifield, and Cook (2009), it is the most commonly appearing color category, about 5% more common than its closest competitor, black, and about twice as common as the third most common category.

How our perception of unique white arises is not understood. But whatever its neural basis, the relevant neural encoding may be influenced by functional (or ecological) constraints: “white” might be elicited by a particular surface or type of spectrum or be constrained by the ensemble statistics of natural scenes. The physiological and functional accounts of white are not mutually exclusive.

Physiological accounts propose that white is perceived if a defined set of stable physiological signals is present. One suggestion has been that perception of white arises from the set of signals that are the equilibrium points of the opponent mechanisms of color vision (Mollon, 2006). There are two possibilities: that white occurs at the equilibrium point of the opponent mechanisms of the retina and the lateral geniculate nucleus, S/(L+M) and L/(L+M) (Derrington, Krauskopf, & Lennie, 1984) or that white occurs at the equilibrium point of the postulated Hering opponent mechanisms, red/green and blue/yellow (Hering, 1878, 1964; Hurvich & Jameson, 1957). Although the latter remains a real possibility, there have been few anatomical and physiological results that suggest a plausible substrate in the visual system for the Hering opponent mechanisms (see General discussion). In favor of the physiological hypothesis, Walraven and Werner (1991) have found that two observers’ unique white settings are stable over a four log unit variation in
luminance. This sets white apart from other colors, which undergo a Bezold-Brücke hue shift and a progressive desaturation (a regression toward white) as intensity increases. Because the Bezold-Brücke effect is thought to arise from the compression of responses in color channels, Walraven and Werner proposed that the absence of a Bezold-Brücke effect for white is evidence that it has a nonpolarizing effect on color channels.

In a functional perspective, the perception of white could arise from something physically present in the environment, a particular type of spectrum, or a class of objects with particular properties. One candidate spectrum is equal-energy white, which has constant energy per unit wavelength across the spectrum. But most observers perceive this spectrum as faintly pinkish or yellowish. Alternatively, white could correspond to surfaces with a particular class of reflectance spectra. Surfaces that reflect all wavelengths equally are usually perceived as white under natural viewing conditions. However, the visual system does not have direct access to spectra but only to triplets of cone excitations. It would have to learn which triplets of cone excitations correspond to white spectra. It could do this perhaps by taking account of the statistics of natural scenes or by attending to the chromatic gradients arising from highlights (Mollon, 2006).

Theories of color constancy have led to explanations of white perception in terms of the information contained in natural scenes. The gray world hypothesis (e.g., Land, 1983; Land & McCann, 1971) proposes that the average chromaticity in any natural scene is gray, so what the visual system must do to maintain color constancy is to calculate the average chromaticity in the scene and set it as the achromatic point. The gray world hypothesis implies a statistical account of white perception: The chromaticity that an observer perceives as white would depend on the average or weighted average chromaticity of a particular scene.

A relaxation of the gray world hypothesis leads to what we will call the “achromatic centroid principle”: that although individual scenes may have a recognizable predominant coloration, the observer may perceive as white the average chromaticity over a large collection of natural scenes that he has encountered in his visual experience. This would not allow the observer to achieve color constancy for any particular scene, but a long-term running average of all natural surfaces may provide a good estimate of the average illuminant, which may be perceived as white.

The “bright is white” hypothesis for color constancy (Land & McCann, 1971) proposes that color constancy is achieved by perceptual anchoring to the brightest surface in a scene, which is perceived as white. But white can be perceived in the presence of brighter stimuli, if the brighter stimuli are perceived as self-luminous (Evans, 1959). The brightest surface in a scene may be perceived as colored whereas a dimmer surface is perceived as achromatic, albeit as gray rather than as white (e.g., Uchikawa, Fukuda, Kitazawa, & MacLeod, 2012).

Rationale

There is great variability in settings of unique white both within a single observer across different trials and between observers. In Experiment 1, we sought to characterize the locus of this variability because it may shed light on the basis of white perception. Specifically, different patterns in variability of white settings favor different theories for the basis of white.

Where would we expect the variability in white settings to lie if white perception arises when the cone-opponent mechanisms are in a null or equilibrium state? The cone-opponent axes (Derrington et al., 1984) have null planes defined by simple ratios of the S, M, and L cone excitations, S/(L+M) and L/(L+M), where the denominator L+M corresponds to luminance (MacLeod & Boynton, 1979). Contrast sensitivity is much greater for L than for S cones at constant luminance (LeGrand, 1949; Rodieck, 1973, chapter 23). Thus, if the differential sensitivity of those mechanisms is what limits the variability in white settings, then these two cardinal axes of color space will be the axes of maximum and minimum variability.

Alternatively, if white corresponds to surfaces that reflect all wavelengths of light, then we might intuitively expect the locus of maximum variability in white settings to follow the variability that exists in the chromaticities of such surfaces in natural scenes (we develop an explicit rationale for this in the Discussion). This variability is largely caused by the different colors of natural illuminants. The locus of natural daylight runs close to the cerulean line, a locus of colors that extends from spectral unique blue at 476 nm to spectral unique yellow at 576 nm (Larimer, Krantz, & Cicerone, 1974; Mollon, 2006; Panorgias, Kulikowski, Parry, McKeefry, & Murray, 2012) and is close to the Planckian locus for black body radiation. The cerulean line is oblique in the MacLeod-Boynton chromaticity diagram, running from the top left quadrant to the bottom right quadrant (Figure 1). Thus, if white perception corresponds to surfaces that reflect all wavelengths of incoming light, we would expect most variability in settings of unique white along a similarly tilted line.

General methods

Stimuli were presented on a Diamond Pro 2070SB CRT monitor (Mitsubishi, Tokyo, Japan) running at
120 Hz. Gamma correction was achieved using a UDT photometer (United Detector Technologies, Hawthorne, CA) and the color calibration was achieved using a Spectrascan PR650 spectroradiometer (Photo Research Inc., Chatsworth, CA). Experiments were run in Matlab R2007b (The MathWorks, Natick, MA) and stimuli created and presented using a vsg2/4 graphics card (Cambridge Research Systems, Rochester, UK). The viewing distance was set, using a fixed seat, at 50 cm. A chinrest was not used. All experiments were conducted in a dark room.

Participants were undergraduate students who participated in exchange for course credit. Different participants took part in each of Experiments 1–4: the numbers are provided in the relevant sections. All participants had normal color vision assessed using the Ishihara plates (1987 edition, presented under a MacBeth Illuminant C). All participants gave written informed consent before taking part. The study adhered to the principles of the Declaration of Helsinki.

Methods

The test stimulus was a disc of 1.9° diameter and luminance of 24 cd/m². In some conditions, the test disk was presented continuously at the right of fixation. In other conditions, we minimized visual fatigue by presenting the test disk successively at two, eight, or 12 positions arranged around a central single-pixel fixation dot, moving it every second. The test disc was presented at an eccentricity of 3.8°, necessary to allow 12 independent positions of the 1.9 test disc. The color of the fixation dot was the same as that of the disc and changed with the disc in order not to supply a reference for the white judgment.

Participants were instructed to fixate the fixation dot and adjust the color of the test disc using a mouse trackball until it appeared to be unique white. Unique white was defined as “a pure white that contains no hint of any other color—that is neither reddish nor greenish nor bluish nor yellowish.” Participants were further instructed to be very particular about their matches and not to accept any settings that were off-white in any direction. Twenty-four subjects made 35 settings each, and nine subjects made 25 settings each: This typically took between 30 min and 1 hr.

As participants manipulated the mouse trackball, the chromaticity of the stimulus was altered in the MacLeod and Boynton (1979) chromaticity diagram, so that a horizontal change in the position of the trackball caused a corresponding change in the L/(L+M) value of the stimulus, and a vertical change in the position of the trackball caused a corresponding change in the S/(L+M) value of the stimulus. Two outlying data points of S/(L+M) > 2.2 do not appear on this plot. (b) For each participant, the mean white setting is indicated by a circle and the axis of maximum variability by a continuous black line. The line subtends one standard deviation on either side of the mean.

Experiment 1: Settings of unique white in a dark surround

In Experiment 1, we measured the variability of settings of unique white made in a context-free environment, where the adjustable stimulus was embedded in a dark surround.

Figure 1. Settings of absolute white from all participants. (a) Each data point represents one setting for one participant. The dashed line is the cerulean line. The dotted line is the locus of CIE daylights. The continuous black line is the orthogonal linear regression through the data. The locations of equal-energy white (EEW), D65, and CIE daylights of particular (labeled) color temperatures are marked. (b) For each participant, the mean white setting is indicated by a circle and the axis of maximum variability by a continuous black line. The line subtends one standard deviation on either side of the mean.
change in the S/(L+M) value. Participants were not able to adjust luminance. White settings were accepted by clicking a mouse button. Before starting the task, participants were introduced to a colored representation of the MacLeod-Boynton chromaticity diagram. They were asked to memorize the directions in which different colors (red, orange, chartreuse, green, teal, blue, and violet) lay. They had a lamp that they could use to refer to the diagram if necessary, but they were instructed not to adjust the stimulus or to accept a setting unless the lamp was off. Our choice to align the X and Y directions of the mouse trackball with the cardinal directions of the MacLeod-Boynton chromaticity diagram rather than with the blue-yellow and red-green Hering axes was a necessary one. If we had chosen the latter, any variability differences that we might measure to be aligned with the Hering axes could be trivially explained by differences in the gain or accuracy of X and Y motor responses.

The reason for varying the number of positions through which the test disc moved was that we were interested in the effect of adaptation on the variability of white settings. In a previous study measuring the possible impact of macular pigment on settings of unique white (Beer, Wortman, Horwits, & MacLeod, 2005), we were surprised at the extent of the variability in settings. One possible reason for the large variance was adaptation. Because the stimuli did not move and always stimulated the same retinal location, adaptation may have desensitized that area of retina to subtle changes in color. Here we investigated the effect of adaptation by manipulating the amount of time each area of the retina was stimulated: In the case in which the test disc moved through two positions, only two areas of the retina were stimulated 50% of the time each, and so forth.

Thirty-three participants took part in Experiment 1.

Results

All settings of unique white are shown in Figure 1a. Each data point represents one setting of white by one participant. The curved dotted line indicates the CIE daylight locus, constructed by estimating the spectra corresponding to different color temperatures by the method given by Wyszecki and Stiles (1982, pp. 145–147). MacLeod-Boynton chromaticity coordinates were calculated for each point along the locus, using the same cone fundamentals (Stockman, MacLeod, & Johnson, 1993) that we used to specify our stimuli. In this version of the MacLeod-Boynton chromaticity diagram that we use throughout, the axes are scaled so that equal-energy white has the coordinates (0.7, 1). The L/(L+M) coordinate of 0.7 arises because for the assumed cone sensitivities, equal energy derives 70% of its luminance (assessed by the CIE 10 degree luminosity function Y10(l)) from the L cones and 30% from the M cones.

The dashed line in Figure 1 is the cerulean line, and the continuous black line shows an orthogonal linear regression through the distribution of settings. For the orthogonal linear regression, the S/(L+M) and L/(L+M) axes of the diagram were first scaled to normalize the overall variance along each axis. This is because the scaling factor for the S cone sensitivity function relative to the sensitivity functions for the L and M cones is otherwise arbitrary, and the scaling of the S cone sensitivity function has an impact on the angle of the resulting regression line. The orthogonal linear regression was transformed back into our MacLeod-Boynton chromaticity coordinates for plotting. The orthogonal linear regression, which is also the axis of maximum variability in the data, lies very close to the cerulean line and to the locus of CIE daylights both in position and in orientation. The match to the daylight locus is particularly close; the slope of the cerulean line is slightly shallower, with 0.82 times the slope.

In Figure 1b we show, for each participant, the mean setting of unique white (open circles) and the axis of maximum variability of settings (solid lines). This axis has again been calculated in a space where the variances along the x- and y-axes are normalized (for each participant) and then transformed back into our standard chromaticity diagram. For all 33 participants, the axis of maximum variability in settings lies along the negative diagonal. In all but a few cases, the angle of this axis is very close to those of the locus of CIE daylights and the cerulean line (also shown). The orientation of the lines in Figure 1b is determined by the ratio of the horizontal and vertical standard deviations of each participant’s settings.

The arithmetic mean setting for unique white across all participants has the coordinates \( L/(L+M) = 0.687 \) and \( S/(L+M) = 1.16 \). This setting is significantly lower in \( L/(L+M) \) (\( t = 12.6, p < 0.001 \)) and higher in \( S/(L+M) \) (\( t = 8.1, p < 0.001 \)) than the coordinates of equal-energy white. The mean setting is closer to the chromaticity of D65 (\( L/(L+M) = 0.690, S/(L+M) = 1.08 \)) but still significantly lower in \( L/(L+M) \) (\( t = 3.17, p = 0.003 \)) and higher in \( S/(L+M) \) (\( t = 4.71, p < 0.001 \)). Equal-energy white and D65 are slightly displaced from the mean white approximately along the cerulean line in the yellow direction. In the distribution of individuals’ mean settings (Figure 1b), there is a significant negative correlation between \( S/(L+M) \) and \( L/(L+M) \) of 0.41 (\( p = 0.02 \)). Thus, the excess variability we observe along the negative diagonal of the MacLeod-Boynton chromaticity diagram is present across individuals as well as within individuals.
To investigate the effect of varying the number of positions that the test disk moved through, and therefore the level of adaptation at stimulated retinal locations, we calculated for each participant the standard deviation of settings of unique white for each condition. There was no significant effect of number of test disk positions on either the mean or the variance of white settings ($F(3, 128) = 0.91, p = 0.44$, for the mean; Levene’s statistic $= 0.82, p = 0.49$, for the variance). However, when the test disk moved, rather than remaining stationary, the task was a much more comfortable one for the participants, because they did not experience Troxler fading of the test disk.

**Interim discussion: Experiment 1**

There is more variability in settings of absolute white along the negative diagonal in the MacLeod-Boynton chromaticity diagram than along any other axis. The location of the line of maximum variability lies very close to the cerulean line and to the locus of CIE natural daylights. The pattern of results we see from all of the data (Figure 1a) is very consistent across the 33 individual participants (Figure 1b).

The blue-yellow direction of maximum variability in white settings has already been discussed by McDermott and Webster (2012) and is evident in several existing data sets including those of Honjo and Nonaka (1970), Webster and Leonard (2008), and Werner and Schefrin (1993). But what might it reveal about the nature of white? One possible cause of the observed distribution of white settings is uncertainty about illumination: An observer may perceive as white any stimulus that could plausibly arise from the product of a flat reflectance spectrum and the spectrum of a natural illuminant. Because a surface with a flat reflectance spectrum in the real world can have any chromaticity depending on the illumination, to set any particular chromaticity as white, observers may have to have a representation of the chromaticity of the illuminant, which may vary between observers or across trials. As the daylight locus of Figure 1 suggests, stimuli that have chromaticities that are the product of a flat reflectance spectrum and a natural illuminant lie close to the blue-yellow cerulean line. The figure does not reveal the extent of the negative diagonal elongation in the distribution of individual participants’ settings, which is discussed below in conjunction with the later experiments.

Experiments 2–4 were designed to address the hypothesis that the increased variability in settings of absolute white along the negative diagonal is traceable to varying color constancy corrections associated with different assumed natural illuminants that are characteristically concentrated along the cerulean line. We took two experimental approaches:

1. **In Experiments 2 and 3**, we manipulated the extent to which the task required absolute or relative color judgments: Variability along the negative diagonal should be greater for absolute than for relative judgments. Relative judgments (discriminating one color from another) do not require an accurate estimation of the illuminant, whereas absolute judgments do, because the chromaticity of the illuminant must be known in order to discover an object’s reflectance spectrum from the product of illumination and reflectance spectrums that is received by the eye.
2. **In Experiment 4**, we compared color discrimination on black and white surrounds. We expected that providing an explicit cue to the illumination in the form of a white surround would reduce variability in the blue-yellow direction.

---

**Experiment 2: Absolute settings versus matching and forced-choice discrimination**

In Experiment 2, we compared settings of white made in a dark surround (absolute judgments) with bipartite matches and four-alternative forced-choice color discrimination (relative judgments).

**Methods**

For the bipartite matching task, the stimulus was a centrally presented vertically divided circular field of diameter $2.1^\circ$ and luminance $24$ cd/m$^2$. One half, selected randomly on each trial, was metameric with equal-energy white, whereas the other half was adjustable. The participant’s task was to alter the chromaticity of the adjustable half of the bipartite field until the two halves appeared to match in color, forming a uniform circle. The method of adjustment was the same as that used in Experiment 1. Participants made as many matches as they could in two blocks of 10 min each.

For the four-alternative forced-choice task, the stimulus was an annulus divided into four quadrants, isoluminant at $24$ cd/m$^2$, presented around a central fixation dot. The outer diameter of the annulus was $2.1^\circ$, and the inner diameter was $0.8^\circ$. There were gaps of $0.15^\circ$ between each segment. On each trial, three segments were metameric with equal-energy white, and the fourth was of a different chromaticity. The
participant’s task was to select the segment that differed in color from the other three. Two interleaved ZEST staircases (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Watson & Pelli, 1983) tracked the participant’s discrimination threshold along lines radiating away from equal-energy white. Discrimination thresholds were measured in eight different color directions in separate randomly ordered blocks. The color directions were increments and decrements along the cardinal axes of the MacLeod-Boynton space (four directions), increments and decrements along a line parallel to the cerulean line and passing through equal-energy white (two directions), and increments and decrements along the reflection of the line parallel to the cerulean line (two directions).

Fifty-one participants took part in the bipartite matching task, and 35 independent participants took part in the four-alternative forced-choice task.

Analysis

To compare the distributions of variability in color judgments between conditions, we fit ellipses to the data, first transforming each subject’s data as for the orthogonal linear regression in Experiment 1 (Figure 1), by normalizing the variances of the data along the two cardinal axes of the MacLeod-Boynton chromaticity diagram. For the absolute judgments from Experiment 1 and for the bipartite matching data, we fit standard deviation ellipses (Yuill, 1971). In the four-alternative forced-choice experiment, threshold chromaticity differences obtained independently for the eight fixed-color directions are represented by eight points appropriately distant from the fixed ellipse center in the chromaticity diagram (Figure 2c). For each color direction, the proportion correct was fit by a Weibull function of the log of the distance in chromaticity from the fixed ellipse center, scaled to range between asymptotes of chance performance (25% correct) for zero chromaticity difference (a log difference of −Infinity), and 99% correct for large differences. We defined as threshold the 81% point on the psychometric function. We then found the free ellipse parameters that minimized the sum of the squared difference between the observed and expected log thresholds.

The best-fitting ellipse was estimated without iteration as follows. The chromaticity coordinates of the eight measured threshold points (here x and y for short) were used to define five row vectors of length 8, one row for each of the five terms in the second-order polynomial defining an ellipse:

\[ c_{20}x^2 + c_{11}xy + c_{02}y^2 + c_{10}x + c_{01}y = 1 \]

The values of the coefficients \( c \) for each of the five terms were then obtained in MATLAB by a right matrix division of an eight-long row of ones by a 5 × 8 matrix containing the experimental values of \( x^2, xy, y^2, x, \) and \( y \). The resulting coefficients define a metric for color difference (relative to the center of the ellipse) that has the value of 1 along the ellipse contour and has computable values for each of the eight measured threshold points. MATLAB chooses the ellipse coefficients to minimize the mean squared deviation of the threshold points from 1. This is approximately equivalent to minimizing the mean squared ratio of measured to predicted threshold or to minimizing the mean squared difference between measured and predicted log threshold along each chosen radius. The resulting fit attaches equal costs to equal fractional errors of prediction. This is appropriate if log threshold, rather than linear threshold, is roughly Gaussian and uniform for the different measured directions in color space.

To quantify the extent to which performance is reduced in a blue-yellow direction, we extracted from the fitted ellipses in the normalized chromaticity diagrams axis ratios: the ratio of the length of the negatively sloping axis to the length of the positively sloping axis. In the normalized MacLeod-Boynton chromaticity diagram, there are only three possible orientations of the fitted ellipse: Its major axis will lie along the negative diagonal or along the positive diagonal, or the fit will be circular. If an axis ratio is greater than 1, then the data show more variability along the negative diagonal than along the positive diagonal. Example data for one participant are shown in Figure 2a to c. The axis ratio has a simple relationship to the correlation between the \( S/(L+M) \) and the \( L/(L+M) \) coordinates of the data points: Denoting that correlation by \( r \), the square of the axis ratio is \( (1 − r)/(1 + r) \).

Results

For bipartite matching, results are based on matches from 51 participants. The mean number of matches made was 52.5 (range = 26–176; \( SD = 27 \)). There was evidence of a speed-accuracy tradeoff in performance: significant correlations between the standard deviation of matches and the number of matches made (\( r = 0.48, p = 3.5 \times 10^{-4} \)) with the standard deviation of matches in \( L/(L+M) \); \( r = 0.37, p = 0.007 \), with the standard deviation of matches in \( S/(L+M) \). However, there was no significant correlation between the number of matches made and the axis ratio of match variability (\( r = −0.13, p = 0.37 \)). Participants showed the same degree of excess variability in the blue-yellow direction.
whether they made a smaller number of more careful matches or a larger number of faster matches.

For four-alternative forced-choice, results are based on 16 threshold estimates (two for each color direction) for each of 35 subjects. Mean axis ratios for all conditions are shown in Figure 2e. They are 2.4 for absolute judgments, 2.4 for bipartite matching, and 2.7 for four-alternative forced choice. There is no significant difference in axis ratio between the three conditions. The excess variability along the negative diagonal seen in the judgments of absolute white (Experiment 1) is equally apparent in the relative judgments required for bipartite matching and four-alternative forced choice in Experiment 2.

Although the axis ratio is impressively invariant, the orientations of the ellipses (prior to normalization) differ somewhat between the different experimental procedures. Accordingly, the major axes are not always precisely aligned with the cerulean line. In our cone chromaticity diagram with equal-energy white at (0.7, 1.0) the cerulean line has a slope of −20.8, implying that a 20% change in S/(L+M) is offset by a change of (1/0.7)% or 1.4% in L/(L+M).

In the unnormalized (upper) plots of Figure 2a, b, and c, the slopes of the ellipse major axes depend mainly on the ratio of the vertical to horizontal standard deviations (included in each panel), because the ellipses were first fit in the normalized space (lower plots) and then transformed back into the original space for plotting. The slopes represented in Figure 2d are those ratios. They determine the normalization factors for the normalized plots that appear in the lower panels of Figure 2.

For absolute judgments (Figure 2a), the mean slope given by the ratio of the standard deviations is slightly steeper than the cerulean line. The forced-choice data imply a slope of −20.8, in excellent agreement with the cerulean line. But the ellipses based on bipartite field matches are substantially steeper. Such variation is not unexpected: For instance, small field tritanopia may have reduced the S cone contrast sensitivity in the bipartite matches relative to the forced-choice thresholds (where the center 0.8° of the field was blanked out). The variances of settings may be influenced by the geometry of the trackball used to make them, a factor that we have not investigated. The forced-choice ellipse is not subject to such biases. Distributions of slopes for all conditions are shown in Figure 2d.

Experiment 3: Effect of delay

In Experiment 3, we investigated the effect of delay, comparing simultaneous judgments to successive judgments. We surmised that a successive judgment would
be more absolute than a simultaneous judgment and more affected by fluctuating color constancy corrections.

**Methods**

Our stimuli were tripartite fields similar to those used in the four-alternative forced-choice task of Experiment 2 but with three segments (upper, lower left, and lower right), rather than four. The segments were isoluminant at 12 cd/m². In different blocks, the randomly located test segment varied in chromaticity from the other two segments either incrementally or decrementally along a line parallel to the cerulean line, passing through D65, or else incrementally or decrementally along the reflection of that line. There were two experimental conditions: a simultaneous condition in which all three segments appeared together for 300 ms and a successive condition in which they appeared sequentially, in a random order, each for 300 ms, with an interstimulus interval of 500 ms. In both cases, the participant’s task was to identify the segment that differed from the other two in chromaticity. Two interleaved ZEST staircases, each of 40 trials, tracked the participant’s discrimination threshold in each block.

To force the participant to use relative rather than absolute judgments in the simultaneous condition, we introduced jitter in the saturation of the two distractor segments, around the reference chromaticity metamer of D65. The jitter was drawn from a normal distribution truncated at three standard deviations from the mean, where the mean jitter was the same size as the current threshold in the relevant direction in color space. The mean jitter therefore tracked the participant’s threshold along with the ZEST staircases. There was no such jitter in the successive condition, with the distractor segments always having the chromaticity of D65.

Fifteen participants took part in the experiment.

**Results**

A repeated-measures analysis of variance revealed a significant main effect of the condition (successive versus simultaneous: $F(1, 111) = 34.2; p = 5.1 \times 10^{-8}$, the successive threshold being higher), and a significant main effect of the line tested (cerulean line versus reflection of cerulean line: $F(1, 111) = 52.3; p = 6.4 \times 10^{-11}$, the cerulean threshold being higher). There were no significant interactions. The effect of polarity was not significant ($F(1, 111) = 0.23; p = 0.63$): Thresholds for increments and decrements were not significantly different. This being the case, we collapsed the data across increments and decrements for each line tested. Figure 3a shows mean thresholds for our collapsed data.

To characterize the difference between the simultaneous and successive conditions, we extracted for each condition the ratio of the thresholds along the cerulean line to the thresholds along the reflection. If successive judgments are more affected by uncertainty about the illumination than simultaneous judgments, we would expect this ratio, which is comparable with the axis ratio extracted from the results of Experiment 2, to be greater for the successive condition than for the simultaneous condition. The average ratio for each condition is shown in Figure 3b. The figure shows that threshold ratios are similar for each condition, and a $T$ test confirmed there was no significant difference ($T = 0.36, p = 0.73$). Similarly in Figure 2, absolute judgments (Figure 2a) show greater variability than side-by-side matching (Figure 2b), yet the axis ratio is preserved. In that sense, the results do not uphold our
conjecture that successive or absolute judgments would add variance specifically along the cerulean line owing to fluctuating color constancy corrections. On the other hand, the preservation of the axis ratio despite the added variance seen in the delayed or absolute judgments implies that the added variance, like the variance of simultaneous judgments, is directed preferentially along the cerulean line rather than being completely isotropic: Memory errors and instantaneous errors have the same characteristic cerulean elongation.

Experiment 4: Effect of stimulus surround

In Experiment 4, we measured the effect of the stimulus surround on bipartite matches and four-alternative forced-choice discrimination. The surround was either gray or black. Our hypothesis was that if the increased variability in color judgments we observe along the cerulean line is caused by uncertainty about the illumination, providing a large area gray surround as a cue to the illumination should lead to a selective reduction in variability along the cerulean line, just as a large surround can stabilize perception of lightness (Li & Gilchrist, 1999). We do not mean by using the term cue to imply that color constancy is driven only by high-level processes: The white background may inform the visual system about the color of the illuminant by fixing its state of adaptation. In either case, the background could stabilize perception by constraining an implicit perceptual estimate of the illuminant.

Methods

The methods for Experiment 4 were the same as for Experiment 2, except for the introduction of the two surround conditions. In both conditions, the surround had a width of 40° and a height of 30°. The gray surround had a chromaticity metameric with equal-energy white and a luminance 72% of that of the stimulus. The black surround had the lowest luminance achievable using the CRT monitor (<0.01 cd/m²). Fifty-one participants took part in the bipartite matching task, and 34 participants took part in the four-alternative forced-choice task.

Results

The results of 51 participants for the bipartite matching task are shown in Figure 4a, with a standard deviation ellipse fit to all the data. Results
averaged across 34 participants for the four-alternative forced-choice task are shown in Figure 4b. In both cases, the upper panels show the discrimination ellipses for black and white surrounds in our version of the MacLeod-Boynton chromaticity diagram, and the lower panels show the ellipses in the normalized space we used to extract axis ratios. As in Experiment 2, the ellipses representing the distribution of bipartite matches have a steeper slope than the cerulean line, but the forced-choice threshold ellipses are well aligned with the cerulean line, with slopes of −20.8 and −22.3, respectively, for dark and white surrounds (Figure 4c).

For each participant in each condition, we calculated the axis ratio. Figure 4d shows mean axis ratios for each condition. For both the bipartite matching task and the forced-choice task, the axis ratios were significantly larger when the surround was black than when the surround was white (t = 4.1, p = 1.5 × 10^{-4}, for bipartite matching; t = 6.7, p = 1.2 × 10^{-7}, for forced choice). Moreover, overall variability was greater with the black surround. Thus, the added variance seen with the black surround is not isotropic but is protracted along the negative diagonal.

**General discussion**

**Absolute white settings**

Results from Experiment 1, and results in the existing literature (e.g., Honjo & Nonaka, 1970; McDermott & Webster, 2012; Webster & Leonard, 2008; Werner & Schefrin, 1993) show that there is more variability in settings of unique white along the negative diagonal of the MacLeod-Boynton chromaticity diagram than along the positive diagonal. Because the locus of maximum variability of settings is close to the locus of CIE natural daylight, we considered the hypothesis that color constancy may be responsible for the greater variability that we observe along the negative diagonal. It is plausible that the observer perceives as white any chromaticity that could arise from a “white” reflective surface under natural illumination. In this scheme, color constancy could be seen as introducing variability or “noise” along the negative diagonal, because the observer is uncertain about what the chromaticity of the illuminant is.

We tested our constancy noise hypothesis in two ways. First, in Experiments 2 and 3, we manipulated the degree to which tasks required relative or absolute color judgments. Second, in Experiment 4, we presented the stimuli on a white surround, providing an explicit cue to the illumination.

**Absolute versus relative judgments**

Experiment 1 required participants to make absolute judgments: They had to adjust the chromaticity of a disc presented in complete darkness until it appeared white. Experiment 2 required relative judgments of two kinds: bipartite matching and forced-choice discrimination.

There seems to be no a priori reason why matching and forced-choice discrimination would require the participant to know the chromaticity of the illumination as they would when required to make an absolute judgment. These two relative tasks simply require that the participant discriminate one color from another. For this, it does not seem important to know what the illumination is. It seems likely that the participant would assume that both sides of the stimulus in the bipartite matching task, and all four segments in the four-alternative forced-choice task, are under the same illumination.

If it is unnecessary to know the illuminant to complete the task, and if uncertainty about the illuminant is causing excess variability of white settings along the negative diagonal in the results of Experiment 1, we would expect relative variability along the negative diagonal to be reduced in Experiment 2. We quantified relative variability by fitting ellipses to the data in a normalized space and taking axis ratios of the lengths of the ellipses’ axes along the negative and positive diagonals. We found no significant differences between axis ratios for absolute white judgments (Experiment 1) and for relative judgments (Experiment 2). In all cases, the axis ratio remained large: The average was 2.4 for judgments of absolute white, 2.4 for bipartite matching, and 2.7 for four-alternative forced choice. Discrimination is about 2.5 times worse along the negative diagonal of the MacLeod-Boynton chromaticity diagram than along the positive diagonal, no matter what the method of measurement.

In Experiment 3, we manipulated in a different way the extent to which the color judgments required were relative. We compared simultaneous judgments, which we assumed to be relative, with successive judgments, which we assumed to be more absolute, because they require participants to store stimulus properties in memory. Consistent with the results of Experiment 2, we found no significant effect of condition (simultaneous versus successive) on the extent of reduced sensitivity along the negative diagonal.

The results of Experiments 2 and 3 go against our hypothesis that uncertainty about the illumination increases variability in white settings specifically in a blue-yellow direction. When the task is relative rather than absolute, knowledge about the color of the illuminant should not affect performance. Yet we still find the same degree of reduction in color discrimina-
tion along the cerulean line in the relative tasks as we find in the absolute tasks.

**Providing a cue to the illuminant**

In Experiment 4, we manipulated the surround. In one condition the surround was dark, and in the other it was metameric with equal-energy white, with a luminance 72% of that of the test stimulus. The rationale was that if uncertainty about the illuminant causes the excess variability in white settings that we observe along the negative diagonal in the MacLeod-Boynton chromaticity diagram, the excess variability should be smaller if a white surround is provided as a cue to the illuminant. For both bipartite matching and for a forced-choice discrimination, we found that ellipse axis ratios were indeed significantly larger when the surround was black than when it was white.

**Constancy noise hypothesis**

The finding that excess variability in a blue-yellow direction was the same for relative judgments as for absolute judgments is unexpected under our constancy noise hypothesis, according to which observers have a noisy representation of the illuminant that fluctuates in chromaticity along the cerulean line. The results of Experiments 2 and 3 are in disagreement with that hypothesis, while those of Experiment 4 provide at best questionable support for it. The dark surround condition of Experiment 4 did exhibit more cerulean elongation than the white surround condition but not necessarily because absolute judgments were involved: The judgments could in principle have been made on a strictly relative basis even in the absence of a fixed surround.

Our findings that are broadly against color constancy being responsible for the skewed distribution of settings of absolute white seem to be against recent conclusions by Pearce, Crichton, Mackiewicz, Finlayson, and Hurlbert (2014) that human color constancy is optimized for blue daylight illuminations. Pearce et al. measured illuminant discrimination, finding that it is poorest in CIE Lu*v* space along a blue locus. The results of our Experiment 2 suggest that poor performance in discrimination along the daylight locus is not specific to illumination changes but occurs equally for stimuli likely to signify reflectance changes. A study that conducts measurements of illumination and reflectance discrimination using a method similar to that of Pearce et al. would be instructive.

The recent viral phenomenon of “the dress” (e.g., McCoy, 2015) has spurred a discussion among vision scientists on the putative special status of the daylight locus. An ambiguous photograph is interpreted by some people as showing a white and gold dress and by others as showing a blue and black dress. As Lafer-Sousa, Hermann, and Conway (2015) and Gegenfurtner, Bloj, and Toscani (2015) pointed out, one of the interesting features of the image is that the pixels lie largely along the daylight locus. We think that the daylight locus may be important in generating the individual differences, not because color constancy is particularly good along that axis or because color discrimination is particularly bad, although the latter may be a contributing factor. More likely, the daylight chromaticities of the pixels and cues to the illuminant in the image offer two different but similarly likely priors for the illuminant: one that it is blue and one that it is yellow.

Our results suggest that constancy fluctuations are not the main source of cerulean elongation. They may nevertheless have a limited role. First, in all conditions involving absolute judgment (most clearly, the results shown in Figures 1 and 2a, but also the delay condition of Experiment 3 and the dark surround condition of Experiment 4), there was more variability in the judgments than in conditions involving relative judgments, yet the cerulean elongation was preserved or increased. It is natural to think of this as a multiplication of all variances by an approximately isotropic factor. Alternatively, these conditions can be viewed as adding additional error variance to the judgments, and on that view, the added variance must have at least the degree of elongation seen in the simultaneous judgments. Second, it is possible that the highly anisotropic distribution of natural illuminants compromises our ability to signal color along the cerulean line, not only when absolute judgments are required, but in general. Even in a single natural scene, different sections may be exposed either to blue illumination (from skylight) or to yellow illumination (from sunlight). The visual system must perform color constancy separately for the differently illuminated regions. It may have learned that along the cerulean line, the correspondence between triplets of cone activations and the colors of reflective surfaces is unreliable. It may therefore settle for coarse discriminations along the cerulean line, and this might affect performance on any task, whether it requires an absolute or a relative judgment. These considerations aside, our results are clearly discouraging for the constancy fluctuation hypothesis and motivate a search for an alternative explanation for cerulean elongation.

**Chromaticity ranges and chromaticity correlation**

Functionally, we might ideally expect the precision of absolute judgments to reflect the distribution of
natural illuminants, whereas the precision of relative judgments would similarly be dictated by the distribution of natural surface reflectances. Our results suggest instead a broadly similar statistical picture for both absolute and relative judgments, both showing a similar cerulean elongation.

As noted, movement away from white along the cerulean line changes S and L cone excitations in opposite directions but more rapidly by an order of magnitude for S than for L cones (MacLeod & von der Twer, 2003), an asymmetry also seen in the fact that S/(L+M) varies enormously across the spectrum (Boynton, 1980; McMahon & MacLeod, 1998; Stockman, MacLeod, & Johnson, 1993). The major axis of the distribution of whites must have a slope similar in magnitude to that of the cerulean line (~20) because the range of unnormalized S/(L+M) values exceeds the range of L/(L+M) values by roughly an order of magnitude. Whether the slope is negative or positive depends on the sign of the correlation between S/(L+M) and L/(L+M). The orientation of the ellipse characterizing a bivariate normal distribution depends only on these factors and not on the magnitude of the correlation between the variables. In its slope, the distribution of whites is appropriately matched to the variation in natural illuminants. But in one respect, the distributions fail to agree. The degree of cerulean elongation (represented equivalently by the correlation coefficient or the axis ratio) is far less than for natural illuminants, which are generally idealized as lying along the single daylight locus shown in Figure 1, despite the relatively minor deviations from that locus that have been statistically characterized (Judd et al., 1964). Analysis of a collection of 337 natural illuminants assembled from the literature by Philipona and O’Regan (2006) reveals a correlation between S/(L+M) and L/(L+M) of ~0.935, with a corresponding axis ratio of 5.5, or roughly 2.5 times the axis ratio in the distribution of whites. For a set of 2,600 natural daylight illuminants by Hernández-Andrés, Romero, Nieves, and Lee (2001), there is a correlation of ~0.988 and an axis ratio of 12.8, 5.3 times the axis ratio of the distribution of whites. The distribution of white chromaticities clearly does not match the correlation characteristic of natural illuminants.

Optimum nonlinear encoding of natural spectra

But first we must ask: Why should the distribution of white settings match the distribution of natural reflectances? The idea of a “match” rests on assumptions that are worth making explicit.

We may think of each visual neuron as dividing color space into slices that elicit reliably distinct response levels. The limited dynamic range of neural firing rates dictates a nonlinear neural response function that must provide good discrimination in the presence of noise contaminating the responses—for example, quantization noise reflecting the inevitable random fluctuations in the firing rates of retinal ganglion cells or central neurons. If neural circuitry is to provide a precise representation of the photoreceptor excitations despite such output noise, the neural stimulus-response function should map the full range of outputs onto the range of the inputs (Laughlin, 1981; MacLeod, 2003; MacLeod & von der Twer, 2003). The range in S/(L+M) that must be accommodated is roughly 10-fold greater than the range in L/(L+M); this asymmetry is not strictly environmental in origin but is mainly traceable to internal features of color processing within the visual system. First, the S cone spectral sensitivity curve is displaced some six times further from the L and M sensitivity curves than those are from each other; and second, S cones generally make no contribution to luminance (Eisner & MacLeod, 1980; Ripamonti, Crowther, & Stockman, 2009), whereas L and M cone excitations are strongly correlated with each other and with luminance.

Other things being equal, a neural signal that must encode a 10-fold larger range of inputs should have correspondingly reduced differential sensitivity. If the design objective is to maximize the mutual information between input and output in the presence of output noise, the slope of the chromatic contrast-response function should be proportional to the probability density of encountered stimuli (Laughlin, 1981). Differential sensitivity will then vary in proportion to the width of the probability density function of the distribution of environmental chromaticities. A some-
what similar principle applies if the aim is to minimize the average error in the noise-contaminated estimate of the input stimulus value (MacLeod, 2003; MacLeod & von der Twer, 2003). Thus, the very unequal sensitivity to chromatic contrast in the two cardinal directions is appropriate for the anisotropic distribution of natural stimuli in cone excitation space.

For optimum nonlinear encoding, however, more than this is needed. Color discrimination should vary inversely with the range of chromatic contrasts in all directions in color space rather than in the two cardinal directions only. This is made difficult by the negative correlation between $S/(L+M)$ and $L/(L+M)$ that is present because the $S$ and $L$ cones favor opposite spectral extremes (McDermott & Webster, 2012). On average, the distribution of natural stimuli is not aligned with the cardinal axes; if the cardinal axes are scaled to make the horizontal and vertical dispersions equal, the distribution is elongated along the negative diagonal. Consequently, no adjustment of relative sensitivity in the two cardinal directions alone can match the probability density of inputs so as to optimize efficiency for all directions in color space.

Although sampling considerations make it difficult to document the generality of the mentioned negative correlation for natural spectra, we can offer examples based on data sets of hyperspectral images made available by Ruderman, Cronin, and Chiao (1998); Párraga, Brelstaff, Troscianko, and Moorehead (1998); Foster, Amano, Nascimento, and Foster (2006); and Chakrabarti and Zickler (2011). We fitted standard deviation ellipses to the distribution of the chromaticities of the pixels in each hyperspectral image. The mean axis ratio of the fitted ellipses was 2.00, and the standard deviation of this ratio over the set of images was 1.01. The averages for each data set are compared in Figure 5a with the axis ratios we have found for our color discrimination measurements.

The distribution of the axis ratios for all the images is shown in Figure 5b. The anisotropy in the distribution of the chromaticities of natural spectra is quite comparable with the anisotropy of the distribution of stimuli accepted as “white.” Although we have tried to provide a rationale for it, we note that this correspondence is by no means inevitable. Other things being equal, there are reasons to expect discrimination step size to covary with the range of stimulus values that must be represented, but if confusions along a particular direction in color space were for some reason particularly undesirable, discrimination in that direction could be made more precise simply by allocating more neural hardware to represent it. On the evidence of Figure 5, however, there is no such preferential treatment of particular directions in color space.

### A modulating influence of luminance

In Figure 5, the axis ratio for the natural scenes of Ruderman et al. (1998) is only slightly greater than 1. This could call into question the generality of the cerulean elongation for natural stimuli. We gained some insight into this by considering the axis ratio (or equivalently, the correlation between the two cardinal chromaticity coordinates) as a function of pixel luminance. We grouped individual pixels from the scenes of Ruderman et al. (1998) and of Párraga et al.
The results shown in Figure 6 provide an alternative noise hypothesis in the results of Experiments 2 and 3. White one, despite a lack of support for our constancy yielded a more pronounced cerulean elongation than a white one, despite a lack of support for our constancy noise hypothesis in the results of Experiments 2 and 3. The results shown in Figure 6 provide an alternative functional rationale for this. Stimuli brighter than the surround will in general be of higher luminance and will come from a distribution with a greater cerulean elongation. Perhaps the incremental spatial contrast of a target viewed in a dark surround causes the target to be treated as if it were of greater than average luminance. This could occur if the variation in the response range with direction in color space is different for on- than for off-center cells.

A Hering basis for color?

Although understandable from a functional perspective, this anisotropy of color discrimination is somewhat unexpected in relation to what is known of the organization of the visual system. If discrimination were limited by noise introduced at the photoreceptor level, or at the level of the retinal ganglion cells and the lateral geniculate where color-opponent neurons encode the cardinal directions independently, cerulean elongation of the distribution of whites could not arise. A substantial component of the variability that limits discrimination must originate in signals that encode noncardinal directions, representing the negative diagonal in the normalized chromaticity diagram with limited precision and the positive diagonal with greater precision. These could be the red-green and blue-yellow color-opponent signals postulated by Hering. Cone-opponent signals present in the lateral geniculate nucleus (LGN) may be transformed into those of the Hering primaries at a cortical level (Abramov & Gordon, 1994; De Valois & De Valois, 1993; Guth, 1991; Wuerger, Atkinson, & Cropper, 2005).

The idea that color might be represented according to the Hering primaries at a cortical level has been suggested several times (Abramov & Gordon, 1994; De Valois & De Valois, 1993), but there is little evidence to support it. Although the Hering primaries have been identified as perceptually special colors (the unique hues), there has been no convincing physiological evidence that they are represented differently from other hues in the cortex. Current physiological studies of how color is cortically encoded generally favor the idea that all directions in color space are represented by individual neurons roughly equally (Horwitz & Hass, 2012; Lennie, Krauskopf, & Sclar, 1990; Xiao, Casti, Xiao, & Kaplan, 2007). Any suggestion that there is a Hering red-green channel that lies orthogonal to the cerulean line would have to account for the fact that unique red and unique green are not collinear but on a dog-leg (Ayama, Nakatsue, & Kaiser, 1987; Bosten & Lawrance-Owen, 2014; Burns, Elsner, Pokorny, & Smith, 1984; Miyahara, 2003).

The cerulean elongation evident in judgments of white does call for a noncardinal representation of color, but it
need not take the particular form of the red-green channel postulated by Hering. Although we believe it is likely to be cortical, Danilova and Mollon (2012a, 2012b, 2014) have raised the possibility that a peripheral channel in which input from the S and M cones is synergistic to that from the L cones may underlie the discrimination minima in L/(L+M) that they find at the cerulean line. Such cells have not been as commonly reported in the retina or the LGN as cells that have synergistic M and S cone inputs, but there have been reports of a minority population (e.g., Tailby, Solomon, & Lennie, 2008; Valberg, Lee, & Tigwell, 1986; see Danilova & Mollon, 2012b, for a review).

Physiological and external (environmental) explanations for the reduction in sensitivity we observe for stimuli along the cerulean line are mutually compatible. Mollon (2006) has suggested that unique blue and yellow receive their distinctive phenomenal character from the special environmental status of natural daylights. Putative Hering opponent systems that are polarized by the unique hues may be tuned to the statistics of natural illuminants or natural spectra.

Environmental statistics and uniform color spaces

In contrast to standard color spaces, CIE uniform color spaces (CIE Lab and CIE Luv) aim to provide a color map in which the distance between pairs of colors corresponds to the difference in their appearance. Do the uniform color spaces predict the excess variability in white settings and increased thresholds we observe along the negative diagonal of the MacLeod-Boynton chromaticity diagram? Similarly to McDermott and Webster (2012), we transformed a circle of color coordinates in CIE Lab space and in CIE Luv space into the MacLeod-Boynton chromaticity diagram. In both cases, the results were approximately ellipses with major axes oriented along the negative diagonal. We transformed these ellipses again into the normalized coordinate frame in which we defined our axis ratios. The axis ratios were 1.67 for the ellipse that originated as a circle in CIE Lab space and 1.28 for the ellipse that originated as a circle in CIE Luv space. These axis ratios fall somewhat short of those observed experimentally. For black surrounds, our axis ratios averaged at about 2.5, but for white surrounds, they were lower at about 1.9. CIE Lab goes most of the way toward predicting the reduction we observe in sensitivity along the negative diagonal of MacLeod Boynton space, when the stimulus surround is white. Because CIE Lab is a relative color space and relies on a reference color, it is designed to predict the color appearance of stimuli presented on a surround, and on the evidence presented here, it does that successfully.

What is white?

To return to the question of what it is that underlies perception of white, it is clear from our results and those of many others (Helson & Michels, 1948; Honjyo & Nonaka, 1970; Hurvich & Jameson, 1951; Jameson & Hurvich, 1951a, 1951b; Lee, Dawson, & Smithson, 2012; McDermott & Webster, 2012; Webster & Leonard, 2008; Werner & Schefrin, 1993; Werner & Walraven, 1982) that white is not a single chromaticity. Different observers and even the same observer on different occasions can perceive a surprisingly large range of chromaticities as white, with this range about 2.5 times greater in the blue-yellow direction than in the red-green direction.

Of the possibilities for white that we raised in the Introduction, our results are against the idea that it arises at the null point of the two cardinal color opponent mechanisms, because in this case variability would be more pronounced in a vertical or horizontal direction in the MacLeod-Boynton chromaticity diagram. If white arises at the null point of color opponent mechanisms, they must be more akin to Hering mechanisms, including one tuned orthogonally to the cerulean line.

The wider tolerance for deviations along the cerulean line (that originate from a predominance of long vs. short wavelength light) than for orthogonal deviations (that are associated with curvature of the spectral energy distribution) is against the idea that white corresponds to a flat spectrum of light. Similarly, the negative results of Experiments 2 and 3 are discouraging for the idea that white perception corresponds to surfaces with flat reflectance spectra: Manipulating the requirement for color constancy has little impact on performance. As noted above, our results are more supportive of the hypothesis that white arises from adaptation by the visual system to the nonuniform color statistics of natural scenes. Color discrimination is poorest in the blue-yellow direction because it is in this direction that the greatest diversity of surface reflectances and illumination spectra lies.

Our results provide evidence that the range of chromaticities perceived as white is determined by calibration to the statistics of the external environment. It is therefore plausible that the average white setting is also determined in this way: the achromatic centroid principle. How does our average white setting (L/(L+M) = 0.687 and S/(L+M) = 1.16) compare with the mean chromaticity of natural scenes? The average L/(L+M) coordinate for 40 hyperspectral images by Foster, Nascimento, and Amano (2004), Nascimento, Ferreira, and Foster (2002), Ruderman et al. (1998), and Parraga et al. (1998) was 0.704 (SD = 0.0064), and the average S/(L+M) coordinate was 0.702 (SD = 0.176). This average lies centrally in the lower half of the MacLeod-Boynton chromaticity diagram and is
yellower than our white settings. However, these hyperspectral images are not randomly sampled and do not contain many areas of blue sky or water. Until we have a good random sample of the range of natural spectra encountered by people who live in environments similar to those of our participants, the achromatic centroid principle remains in contention.

All of the theoretical possibilities we have considered imply that physiological processes are conditioned in one way or another by environmental statistics. This dependence could be created either by learning during individual development or by genetic internalization of environmental regularities during evolution. Sensitivity can be modified selectively for different directions in color space in short-term experiments (Webster & Mollon, 1991; Werner, Sharpe, & Zrenner, 2000), and long-term adaptation effects during individual ontogeny can shift the white point as changes in the lens shift the spectrum of light reaching the retina. Settings of unique hues are stable throughout the life span as the lens yellows with age (Schefrin & Werner, 1990; Werner & Schefrin, 1993), but when cataracts are removed, unique white initially shifts in a yellow direction and then gradually returns to its original position (Delahunt, Webster, Ma, & Werner, 2004).

Whether the adaptation involved is evolutionary or developmental, it implicates cortical processes, because the encoding of color in retina and thalamus represents the cardinal directions, \( S/(L+M) \) and \( L/(L+M) \), independently, whereas cerulean elongation requires diagonally tuned mechanisms that presumably originate more centrally. Here we see that what might seem the most primitive possible visual judgment—the discrimination of colors from white—depends critically on cortical processing.

**Keywords:** unique hues, unique white, unique yellow, unique blue, cerulean line, color appearance, color discrimination

## Acknowledgments

We would like to give special thanks to Alexandra Boehm, Natalie Granucci, and Sarah-Nicole Bostan for assistance with running the experiments. The work was supported by National Institutes of Health grant EY01711 (DM).

Commercial relationships: none.
Corresponding author: Jenny M. Bosten.
Email: J.Bosten@sussex.ac.uk.
Address: Department of Psychology, University of California, San Diego, La Jolla, CA, USA.

## References


Li, X., & Gilchrist, A. (1999). Relative area and relative


Walraven, J., & Werner, J. S. (1991). The invariance of


