Does human color constancy incorporate the statistical regularity of natural daylight?

Peter B. Delahunt Department of Ophthalmology and Section of Neurobiology, Physiology and Behavior, University of California Davis, CA, USA

David H. Brainard Department of Psychology, University of Pennsylvania Philadelphia, PA, USA

The chromaticities of natural daylights cluster around the blackbody locus. We investigated whether the mechanisms that mediate human color constancy embody this statistical regularity of the natural environment, so that constancy is best when the illuminant change is one likely to occur. Observers viewed scenes displayed on a CRT-based stereoscope and adjusted a test patch embedded in the scene until it appeared achromatic. Scenes were rendered using physics-based graphics software (RADIANCE) coupled with custom extensions that ensured colorimetric accuracy. Across conditions, both the simulated illuminant and the simulated reflectance of scene objects were varied. Achromatic settings from paired conditions were used to compute a constancy index (CI) that characterizes the stability of object appearance across the two illuminants of the pair. Constancy indices were measured for four illuminant changes from a Neutral illuminant (CIE D65). Two of these changes (Blue and Yellow) were consistent with the statistics of daylight, whereas two (Green and Red) were not. The results indicate that constancy was least across the Red change, as one would expect for the statistics of natural daylight. Constancy for the Green direction, however, exceeded that for the Yellow illuminant change and was comparable to that for the Blue. This result is difficult to reconcile with the hypothesis that mechanisms of human constancy incorporate the statistics of daylights. Some possible reasons for the discrepancy are discussed.

Keywords: color constancy, color appearance

Introduction

The light reflected from an object depends as much on the incident illumination as it does on the object's surface reflectance. Nonetheless, object color appearance is often quite stable across changes in illumination, a phenomenon called *color constancy* (e.g., Kaiser & Boynton, 1996; Brainard, 2003). Indeed, without such stability, it would not be possible to refer to objects as having a well-defined color.

How and under what conditions the visual system achieves color constancy remains mysterious. An important line of research starts with consideration of the computational problem that must be solved by any visual system designed to achieve constancy (see Hurlbert, 1998; Maloney, 1999; Brainard, Kraft, & Longère, 2003). This problem is easily cast, at least for a simplified imaging model.[1] An illuminant is characterized by its *spectral power distribution* $E(\lambda)$. This function yields the power of the incident light at each wavelength in the visible spectrum. An object's surface reflectance is characterized by its *surface reflectance function* $S(\lambda)$. This specifies the fraction of incident light that is reflected at each wavelength. The *color signal* $C(\lambda)$ reflected from the object is obtained as the wavelength-by-wavelength product of the spectral power distribution and surface reflectance function:

$$C(\lambda) = E(\lambda)\, S(\lambda)\,. \qquad (1)$$

Equation 1 makes explicit that the color signal confounds illuminant and reflectance properties. To achieve constancy, the visual system must process the image data to produce a perceptual representation that depends only on reflectance.

Equation 1 provides an *imaging model*, albeit a highly simplified one. The quantities on the right-hand-side of the equation describe the physical properties of a scene. The quantity on the left-hand-side describes the image data available to the visual system. Thus the equation allows calculation of the image data from the scene description.

One way the visual system might attempt to achieve constancy is to invert the imaging model and estimate the object reflectance function. Equation 1 makes clear that this is an *underdetermined inverse problem*: multiple pairs of illuminant spectral power distribution and surface reflectance function can generate the identical color signal. If there are no constraints on the illuminant spectral power distributions and surface reflectance functions that the visual system might encounter, any procedure for inverting Equation 1 will often generate erroneous estimates. If, however, only a restricted range of illuminants and object surfaces are encountered by a visual system, then it is possible to use knowledge of the restricted range to develop sensible estimation procedures. This general observation underlies all modern attempts to solve the computational problem of color constancy (e.g., Buchsbaum, 1980; Maloney & Wandell, 1986; Lee, 1986; D'Zmura & Lennie,

doi:10.1167/4.2.1 Received May 21, 2003; published February 18, 2004 ISSN 1534-7362 © 2004 ARVO

1986; Forsyth, 1990; Funt, Drew, & Ho, 1991; Brainard, Wandell & Cowan, 1989; Trussell & Vrhel, 1991; D'Zmura & Iverson, 1993; D'Zmura, Iverson, & Singer, 1995; Brainard & Freeman, 1997; Finlayson, Hubel, & Hordley, 1997; see Hurlbert, 1998; Maloney, 1999).

The various computational approaches differ in how they elaborate Equation 1 into a more realistic imaging model, and in what constraints they assume about illuminants and surfaces. Most algorithms, however, attempt to incorporate known regularities in the illuminants and surfaces that occur in natural viewing. Figure 1 plots the CIE u'v' chromaticities recently measured for 10,760 natural daylights by Jeffrey DiCarlo and Brian Wandell (DiCarlo & Wandell, 2000). These measurements were made from a rooftop at Stanford University at 1-min intervals from dawn to dusk over a 20-day period in January/February 2000. It is clear from the figure that there is considerable regularity in the locations of the daylight chromaticities. Rather than being distributed uniformly throughout the diagram, the measured chromaticities cluster along a curve, sometimes referred to as the daylight locus. The daylight locus is close to the blackbody locus, which shows the chromaticities of blackbody radiators as a function of color temperature (Wyszecki & Stiles, 1982). Similar regularity is seen in other reported daylight measurements (Judd, MacAdam, & Wyszecki, 1964).

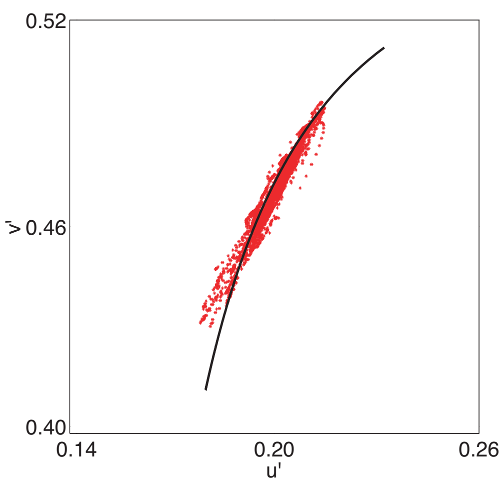

Figure 1. The plot shows the CIE u'v' chromaticity coordinates measured for 10,760 natural daylights. The black curve shows the blackbody locus.

Several authors have shown how the daylight structure revealed by Figure 1 may be exploited by color constancy algorithms (D'Zmura et al., 1995; Brainard & Freeman, 1997; DiCarlo & Wandell, 2000). These algorithms have the feature that constancy will be best across illuminants typical of the daylight locus. If the human visual system is designed to achieve constancy, one might expect that it too takes advantage of the statistical regularities in natural daylights (Shepard, 1992).

Previous comparisons of color constancy across various illumination changes have produced mixed results (also see Discussion). Using real surfaces and illuminations, Brainard (1998) found no advantage for illumination changes along the blackbody locus compared to those off it. Using computer-based stimuli and Mondrian-like patterns, Ruttiger et al. (2001) found less constancy for natural daylight illuminant changes than for red/green illuminant changes. Foster, Amano, and Nascimento (2003) found a similar result, again using Mondrian-like stimuli, and suggested that the lower constancy along the daylight locus might be due to reduced inputs from the S-cones to the constancy mechanism.

The goal of our study was to further investigate how constancy varies with the direction of the illumination change. The experiments measured human color constancy across four illumination change directions. Two of these changes were consistent with changes of natural daylight, while two were not.

Methods

Overview

Consider an object that appears achromatic when $E_1(\lambda)$ is the spectral power distribution of the illuminant. From Equation 1, the color signal reflected from this object will be $C_1(\lambda) = E_1(\lambda)S(\lambda)$, where $S(\lambda)$ is the surface reflectance function of the object. When the same object is illuminated by $E_2(\lambda)$, the reflected color signal will be $C_2(\lambda) = E_2(\lambda)S(\lambda)$. For a color constant visual system, the object should continue to appear achromatic after the change of illuminant. In terms of the color signal, $C_1(\lambda)$ should appear achromatic when the illuminant is $E_1(\lambda)$, whereas $C_2(\lambda)$ should appear achromatic when the illuminant is $E_2(\lambda)$. Our experimental strategy builds on this observation.

A computer-controlled stereo display was used to present synthetic images to observers. Each stereo image pair was generated from a scene description using computer graphics techniques. The scenes consisted of a collection of matte objects, illuminated by a single light source. An example stereo image pair is shown in Figure 2.

Observers viewed a test patch that was embedded in one of the stereo image pairs (see Figure 2). The observers' task was to adjust the test patch until it appeared achromatic. Observers were instructed not to match the test patch with any other object in the scene.[2] Across conditions, the illuminant used to generate the image pair was varied, so that the data consist of the color signals that appeared achromatic when the test patch was viewed in the context of differently illuminated scenes. Comparison of the achromatic adjustments with predictions derived for a color constant visual system leads to a quantitative assessment of constancy.

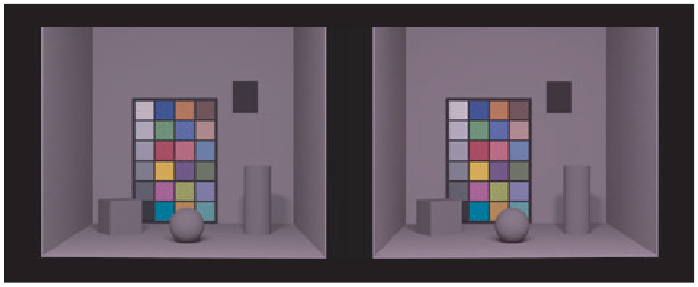

Figure 2. Example of the stereo image pairs used in the experiments. Each image in the pair was synthesized from a three-dimensional scene description using the RADIANCE rendering software (Larson & Shakespeare, 1998). Scene objects were Lambertian. There was a single light source that produced moderately diffuse illumination. For the image pair shown, the Neutral experimental illuminant (see below) was used. The test patch is shown as the dark square toward the upper right of the images. To simulate binocular disparity, the left- and right-eye images were rendered for different viewpoints. The stereo pair in the figure is arranged so that it can be cross-fused.

Achromatic adjustment has been widely used to study color appearance and color constancy (Helson & Michels, 1948; Werner & Walraven, 1982; Fairchild, 1990; Chichilnisky & Wandell, 1996; Bauml, 1994; Brainard, 1998; Kraft & Brainard, 1999; Yang & Maloney, 2001). Speigle and Brainard (1999) showed that measurements of what object appears achromatic under different illuminants may be used to predict how the appearance of other-colored objects will be affected across the same illumination changes.

The following sections provide a basic description of the experimental methods. Methodological details unlikely to be of interest to the casual reader are provided in Appendix A and thorough supplementary material is available by clicking here (supplementary material).

Stimuli

Image synthesis method

The images were synthesized from scene descriptions using the RADIANCE software package (Larson & Shakespeare, 1998). This software uses ray tracing to simulate the flow of light from its source through the scene, and it is intended to produce accurate images. Previous authors have used RADIANCE in psychophysical studies (Yang & Maloney, 2001; Langer & Bülthof, 2000). Although the RADIANCE software provides accurate simulation of light propagation within each color band, the RGB color model is too coarse to provide accurate simulation of the spectral interaction between lights, surfaces, and the human visual system. To remedy this, we wrote custom software that allowed us to extend the color model so that we specified the full spectral power distribution $E(\lambda)$ of each light source and full spectral reflectance function $S(\lambda)$ of each object.

We first specified the spectrum for each of the desired illuminants and surfaces for the scene. Each spectrum was specified by 31 sample values, with samples at 10 nm intervals between 400 and 700 nm. Using a procedure described in Appendix A, 31 simulated monochromatic images were produced and used to compute the excitations of the human L-, M-, and S-cones at each image location, using the Smith-Pokorny estimates of the cone spectral sensitivities (Smith & Pokorny, 1975; tabulated in DeMarco, Pokorny & Smith, 1992). Each cone sensitivity was normalized to a maximum value of 1.

Scene dimensions and content

Figure 3 illustrates the dimensions of the rendering space used for the experiments reported here. The space had dimensions 20 in. (width) × 20 in. (height) × 36 in. (length). The length of 36 in. was the distance from the viewing position to the back of the scene, and in the experiments, the monitors were placed at a viewing distance of 36 in. from the observer.

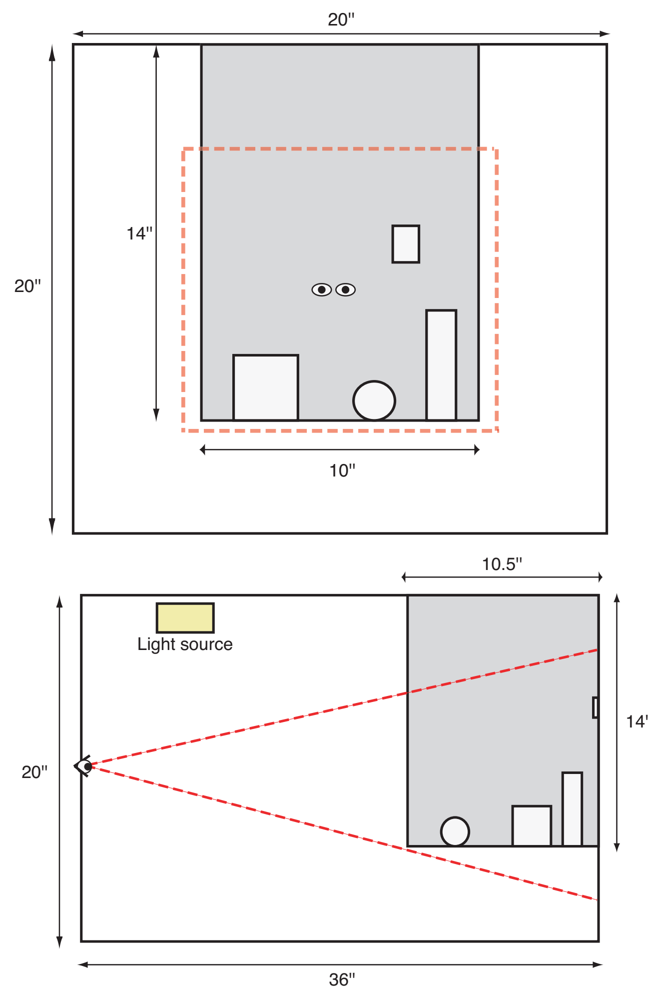

Figure 3. The dimensions of the RADIANCE rendering space used. The top panel shows the view from the observer position. The bottom panel shows the side view. The red dashed lines indicate the approximate area visible to the observer.

The following four simulated objects were placed in the scene: a Macbeth color checker chart, a box, a sphere and a cylinder. More details on the scene dimensions and content can be found in Appendix A.

Illuminants

Nine illuminants were used in the experiments. Their chromaticities are plotted in Figure 4, and their chromaticities and luminances are provided in Table 1. The Neutral illuminant was CIE daylight D65 (CIE, 1986), scaled so that the luminance reflected from a perfect diffuser would be 25 cd/m^2. (In this work, any reference to the luminance of an illuminant refers to the luminance that would be reflected from a perfect diffuser.) CIE daylight D65 has a spectrum corresponding to a typical mixture of direct sunlight and scattered skylight (Wyszecki & Stiles, 1982).

The chromaticities of the Blue_60, Green_60, Yellow_60, and Red_60 illuminants are shown in Figure 4 (solid circles). The Blue_60 and Yellow_60 illuminants were typical of natural daylights, whereas the Red_60 and Green_60 illuminants had chromaticities far from the blackbody/daylight locus. These illuminants had a CIELUV ΔE^* distance from the Neutral illuminant of 60 units. The chromaticities of the Blue_30, Green_30, Yellow_30, and Red_30 illuminants are also shown in Figure 4 (solid triangles). These four illuminants had a CIELUV ΔE^* distance of 30 units from the Neutral illuminant.

Because the Neutral, Blue_60, Blue_30, Yellow_60, and Yellow_30 illuminants are typical of measured daylights, the discussion in the introduction leads to the prediction that constancy would be relatively good across changes between these illuminants. Similarly, because the Green_60, Green_30, Red_60, and Red_30 illuminants are highly atypical of measured daylights, we would expect that constancy would be relatively poor when the illumination changes between the Neutral illuminant and one of these.

Color constancy is by definition a relative phenomenon, because one can ask only about constancy of appearance across some change in the scene. We define the scene with the Neutral illuminant (D65) as the standard scene and assess constancy with respect to changes from this scene. The standard scene is shown in Figure 2.

The distance between the chromaticity of the Neutral illuminant and that of the eight other illuminants was measured in the u*v* chromaticity plane of the CIE 1976 CIELUV uniform color space (CIE, 1986). The rationale for this choice of color space was to equalize as much as possible the perceptual size of the illuminant changes. The distance between the Neutral illuminants and other illuminants was either 30 or 60 ΔE^* units. To calculate illuminant u*v* coordinates requires the specification of a white point. The white point we used in the calculation had CIE xy chromaticity (0.31, 0.33) and luminance 25 cd/m^2. Figure 4 shows the chromaticities of the illuminants in the CIE u'v' (not u*v*) chromaticity diagram. In this representation, the distance between the Neutral illuminant and the other four illuminants is also close to uniform. The advantage of the u'v' representation is that it does not depend on the choice of a white point.

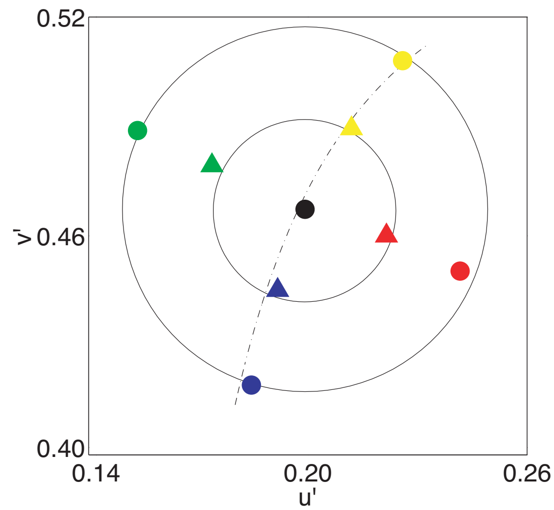

Figure 4. The experimental illuminant chromaticities are illustrated in CIE u'v' coordinates. The closed circles show the illuminants at the perceptual distance of 60 ΔE* units and the triangles show the illuminants at the perceptual distance of 30 ΔE* units. The symbols are color coded so that the Blue illuminants are shown in blue, etc. Both sets of illuminants (60 ΔE* and 30 ΔE*) are approximately equally distant from the Neutral illuminant in the u'v' representation. The blackbody locus is shown by the dashed black curve.

Illuminant	Distance	CIE u'	CIE v'	Lum (cd/m^2)
'Neutral'	None	0.199	0.467	24.4
'Blue_60'	60 ΔE*	0.185	0.419	24.9
'Yellow_60'	60 ΔE*	0.226	0.508	24.9
'Red_60'	60 ΔE*	0.242	0.450	25.3
'Green_60'	60 ΔE*	0.153	0.489	21.9
'Blue_30'	30 ΔE*	0.192	0.445	24.8
'Yellow_30'	30 ΔE*	0.212	0.489	25.2
'Red_30'	30 ΔE*	0.221	0.460	24.7
'Green_30'	30 ΔE*	0.174	0.479	23.9

Table 1. The table provides properties of the nine experimental illuminants. Distance refers to distance from the Neutral illuminant in CIELAB ΔE* units. The indicated luminance is the luminance reflected when the illuminant reflects from a perfect diffuser located on the back wall of the simulated scenes.

Surfaces

All surface reflectance spectra were approximated by a three-dimensional linear model derived from a set of measurements of Munsell papers reported by Nickerson (1957). The reflectance spectra and basis functions are available as part of the supplementary material for this study.

In Experiment 1, the same surfaces were simulated under each illuminant. Here the spectrum of the background surface was chosen so that the chromaticity obtained when the Neutral illuminant reflected from it was that of equal energy white. This same reflectance spectrum was used for the cube, sphere, and cylinder. The light emitted from the monitor from a region on the back wall adjacent to the test patch position was measured after each experimental run. Table 2 provides the mean chromaticities and luminances of these measurements for both Experiment 1 and Experiment 2.

Illuminant	Back wall	CIE u'	CIE v'	Lum (cd/m^2)
'Neutral'	Normal	0.211	0.470	4.89
'Blue_60'	Normal	0.194	0.423	4.95
'Yellow_60'	Normal	0.241	0.509	5.06
'Red_60'	Normal	0.258	0.455	5.20
'Green_60'	Normal	0.160	0.490	4.28
'Blue_60'	Equated	0.211	0.467	4.99
'Yellow_60'	Equated	0.211	0.473	4.83
'Red_60'	Equated	0.210	0.467	5.17
'Green_60'	Equated	0.208	0.472	4.29

Table 2. Results of measurements of the back wall region of the experimental images are provided for all images used in Experiment 1 (labeled "Normal") and Experiment 2 (labeled "Normal" for the Neutral illuminant and labeled "Equated" for the other four illuminants).

The other object in the scene is a simulated Macbeth Color Checker Chart. Reflectance spectra of the patches were obtained from measurements of such a chart made in our laboratory.

In Experiment 2, a different surface was simulated on the back wall under each illuminant. The spectrum of this surface was chosen to hold the light reflected to the observer from the back wall constant across the five illuminants. Under each illuminant, the spectrum was chosen so that the reflected light had (approximately) the chromaticity of equal energy white (CIE u'v' chromaticity of 0.210, 0.471) with a luminance value of approximately 5 cd/m^2 (see Table 2).

Experimental procedure

Observers adjusted the chromaticity of a test patch embedded in the back wall of the simulated scenes until it appeared achromatic (removing all traces of blue, yellow, red, and green). During an adjustment, the luminance of the test patch was held fixed. Observers controlled the chromaticity of the test patch by pushing buttons on a game controller. The button presses changed the CIELAB a* and b* chromaticity coordinates of the test in equal steps. Observers were also able to toggle between three adjustment step sizes by pressing a separate button.

In each session, observers made adjustments for test patches embedded in a single stereo image pair. At the start of the session, the observer adapted to the experimental images for a period of 1 min before making any adjustments. The presentation order of the images across sessions was randomized for each observer. Observers ran in two practice sessions of the experiment before actual data collection began.

Within each session, four different test patch luminance values were used. Two of these were below the luminance of the local surround of the test patch, and two were above. The local surround was the area immediately surrounding the test patch and had a luminance value of approximately 5 cd/m^2. The test patch luminance values were approximately 2.5, 4.0, 6.0, and 8.5 cd/m^2. Each test patch luminance was presented four times making a total of 16 settings per session. One session was typically run per observer per condition.

To assess the reliability of data from a single session, a second session was run. This was done approximately 2 months after the main experiments were completed for a subset of conditions (see Appendix B).

The starting chromaticity of the test patch can affect observers' achromatic settings (Brainard, 1998). In the experiments reported here, an adaptive starting rule was used. At the start of each adjustment, the initial values for a* and b* were chosen by uniform random draw from the range (−25, 25). The conversion from CIELAB coordinates is governed by the choice of a reference white. For the first setting, the reference white used to convert the randomly chosen (a*, b*) starting point had CIE xy chromaticity (0.318, .334) and luminance 17.3 cd/m^2. For each subsequent setting, the reference white was based on a running average of the previous settings in the session.

Observers

Two male and five females were used as observers. The age range was 19 to 37 years. All were color normal as assessed by the Ishihara (1997) plates. Macular stereopsis and far-point acuity was tested using a Keystone orthoscope. All had 20/20 corrected vision or better and normal stereopsis except for one observer (KCC) who was stereoblind and had 20/30 acuity in one eye. The observers were naïve as to the purpose of the experiment except for PBD (one of the authors).

Results

Experiment 1

Figure 5 shows the five experimental images used in Experiment 1. The surfaces in the simulated scene are the same for all the images, while the simulated illuminant differs. The Neutral, Blue_60, Yellow_60, Red_60, and Green_60 illuminants were used (see Table 1). For each observer, the data consist of the achromatic settings made

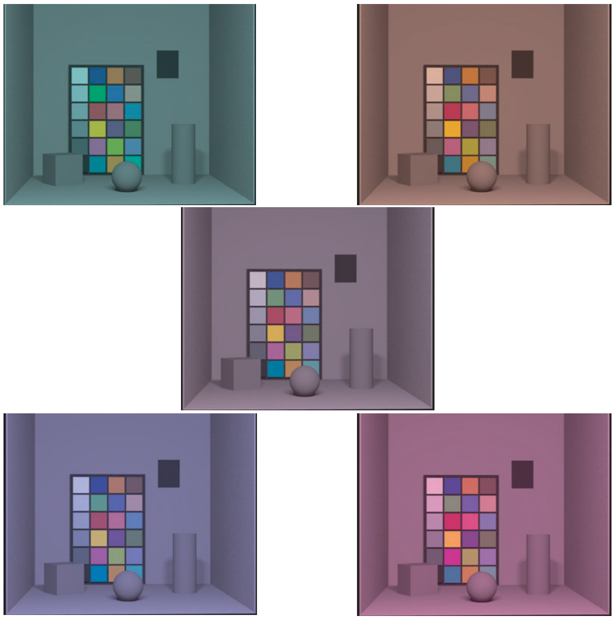

Figure 5. Images used in Experiment 1. One member of each stereo pair is shown. Top row: Green_60 illuminant and Yellow_60 illuminant. Middle row: Neutral illuminant. Bottom row: Blue_60 illuminant and Red_60 illuminant.

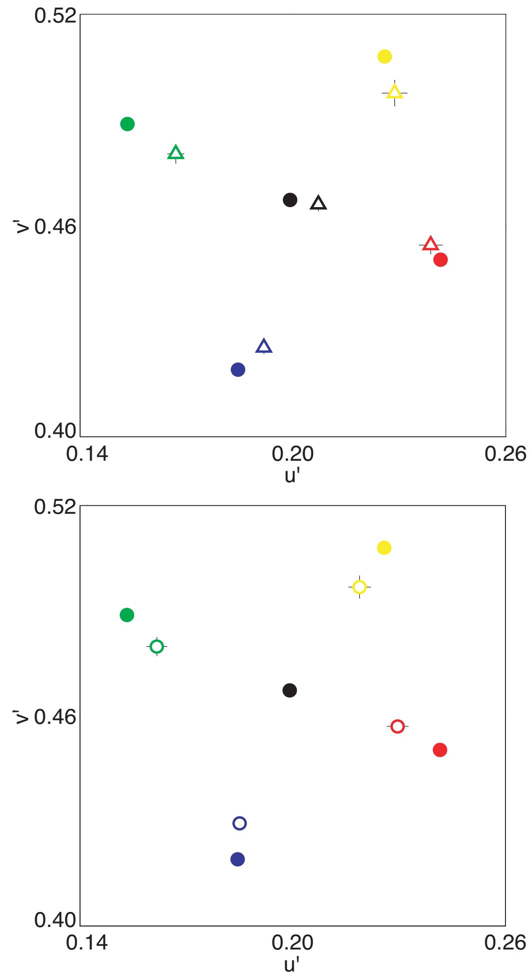

Figure 6. Settings from Experiment 1. Top panel: Achromatic chromaticities averaged over data from seven observers (open triangles) and chromaticities of corresponding experimental illuminants (solid circles). Bottom panel: Equivalent illuminants derived from the achromatic chromaticities (open circles) and chromaticities of corresponding experimental illuminants (solid circles). Where visible, error bars show +/– 1 SEM.

at each luminance for each experimental image. For a single observer, the data for each image may be summarized by the *achromatic chromaticity*, obtained by averaging the u'v' chromaticities of the settings made at the different luminance levels. The difference between settings for luminance values above the value of the surround (increments) and those below the value of the surround (decrements) is discussed in Appendix C.

The top panel in Figure 6 shows the group data. Each plotted chromaticity (open triangles) is the average of the achromatic chromaticities for the seven observers. The color of the plotted points indicates the corresponding experimental illuminant. The chromaticities of the experimental illuminants are also plotted (solid circles). In general, the achromatic settings lie in the vicinity of their corresponding illuminants.

We wish to interpret the achromatic chromaticities in terms of their relation to color constancy. If the visual system made no adjustment at all to the changes in illuminant across our five experimental images, then the relation between the color signal reaching the eye and color appearance should be the same for test patches situated in all five images. Thus for a visual system with no color constancy at all, the five measured achromatic loci should superimpose. Clearly this is not the case for our data.

To understand how achromatic points should vary with the illuminant for a visual system that does have constancy, it is helpful to consider a surface that reflects light equally at all wavelengths in the visible spectrum. Such surfaces are called *spectrally non-selective*, and under a wide range of conditions they appear achromatic or nearly so. In addition, spectrally non-selective surfaces have the physical property that the light reflected from them has the same chromaticity as the illuminant impinging upon them.

Consider the hypothetical case in which (1) the visual system was perfectly color constant and (2) the surface that appeared exactly achromatic was spectrally non-selective. In this case, the measured achromatic chromaticities for each of our experimental images would coincide perfectly with the chromaticities of the simulated illuminants. This is also not the case for the data shown in Figure 6 - each achromatic chromaticity is offset from its corresponding illuminant chromaticity.

How should these offsets be interpreted? Brainard (1998) reported a procedure for recentering a set of achromatic data so that the achromatic chromaticity measured under one chosen *reference illuminant* coincides exactly with the chromaticity of that illuminant.[3] The recentering procedure can be thought of as a model-based prediction of how the entire data set would have looked, had the surface that appeared achromatic under the reference illuminant been non-selective. The bottom panel of Figure 6 shows the result of applying the recentering procedure to the data shown in the top panel, when the Neutral illuminant was selected as the reference illuminant. The recentered achromatic chromaticities are called the equivalent illuminants, one corresponding to each experimental illuminant.

The equivalent illuminant plot is more easily interpreted in terms of constancy. If the equivalent illuminant chromaticity coincides with that of the reference illuminant, no constancy is indicated. If the equivalent illuminant chromaticity coincides with the chromaticity of its corresponding experimental illuminant, perfect constancy is indicated. The bottom panel of Figure 6 shows that the equivalent illuminants from Experiment 1 plot near to the chords connecting the reference illuminant to the experimental illuminants. How far along the chord each equivalent illuminant plots can be taken as a measure of the degree of constancy shown.

We used a constancy index (*CI*) to quantify the degree of constancy with respect to shifts in illumination between two illuminants (see Brainard & Wandell, 1991; Arend, Reeves, Schirillo, & Goldstein, 1991; Brainard, Brunt, & Speigle, 1997; Brainard, 1998). The formula for the constancy index is

$$CI = 1 - [|\,\mathbf{e}_2 - \mathbf{e}_{eq}| \,/\, |\,\mathbf{e}_2 - \mathbf{e}_1\,|] \quad (2)$$

where $\mathbf{e}_1$ is a two-dimensional vector specifying the chromaticity of the reference illuminant, $\mathbf{e}_2$ is a vector representing the chromaticity of the experimental illuminant, and $\mathbf{e}_{eq}$ is a vector representing the chromaticity of the equivalent illuminant. When defined in this way, a constancy index of 1 indicates perfect constancy, whereas a constancy index of 0 indicates no constancy.

We focused on illumination changes between the Neutral illuminant and each of the other four illuminants. We refer to these as the experimental illuminants. For each of these illuminant pairs, we calculated the constancy index first with the Neutral illuminant playing the role of the reference illuminant and then with the experimental illuminant playing the role of the reference illuminant. We report the mean of these two calculations as the constancy index for the illuminant pair.

Figure 7 shows the average constancy indices obtained for Experiment 1 for each of the four illumination changes studied. The indices are all fairly high, ranging between 0.67 and 0.81. We refer to changes between the Neutral illuminant and the four experimental illuminants as the Blue_60, Yellow_60, Red_60, and Green_60 illuminant changes. A one-way within-observer ANOVA indicated that the differences across illuminant changes were not statistically significant ($F(3, 18) = 2.26$, $p = .12$).

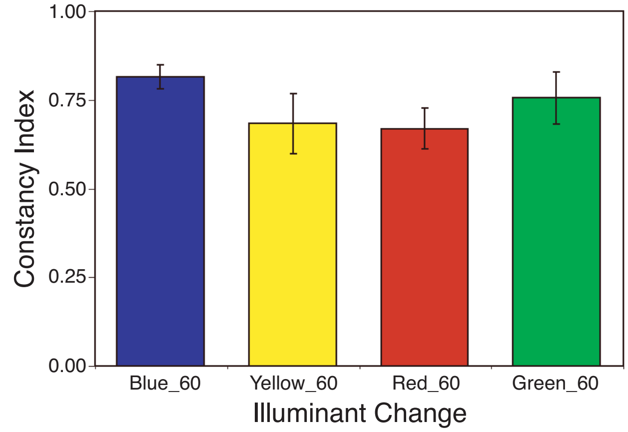

Figure 7. The mean constancy indices obtained in Experiment 1 are shown for each illuminant change . The error bars show +/– 1 SEM.

Experiment 2

In Experiment 1, the *CI*s did not differ significantly with the color direction of the illuminant change. One possibility is that the indices measured in Experiment 1 are subject to a ceiling effect. Because the surfaces in the simulated scenes remained constant across the different illuminants, the images used in Experiment 1 contained many valid cues about the illuminant change (see Kraft et al.,2002; also Yang & Maloney, 2001). It could be that because the quality of information about the illuminant changes across the images used in Experiment 1 was high, the effect of prior information about the distribution of illuminants was masked. We wondered whether an experiment with stimuli that led to lower constancy overall might better reveal an effect of prior information.

Kraft and Brainard (1999; also Kraft et al., 2002) were able to reduce the level of constancy across changes in illumination by reducing the validity of potential cues to the illuminant. In their control condition, Kraft and Brainard (1999) used a design analogous to our Experiment 1. In their “Local Surround” condition, the background surface of their experimental scene was changed for each of the illuminants so that the light reaching the observer in each case was the same. By equating the background, they reduced the mean *CI* from its control condition value (0.83) to 0.53. Here a similar method was used to reduce the overall level of constancy. Experiment 2 was a replication of Experiment 1 with one important change: for each experimental illuminant, the simulated reflectance of the background surface in the scene changed so as to equate the chromaticity and luminance of the light reaching the observer from that region of the image (see Table 2 above.) Following Kraft et al. (2002), we refer to the conditions of Experiment 2 as *invalid-cue* conditions. This term indicates that in Experiment 2 some of the potential cues to the illuminant are invalid in scenes rendered under the experimental illuminants. These conditions may be contrasted to those of Experiment 1, which we refer to as *valid-cue* conditions. Figure 8 shows the images used in Experiment 2. The

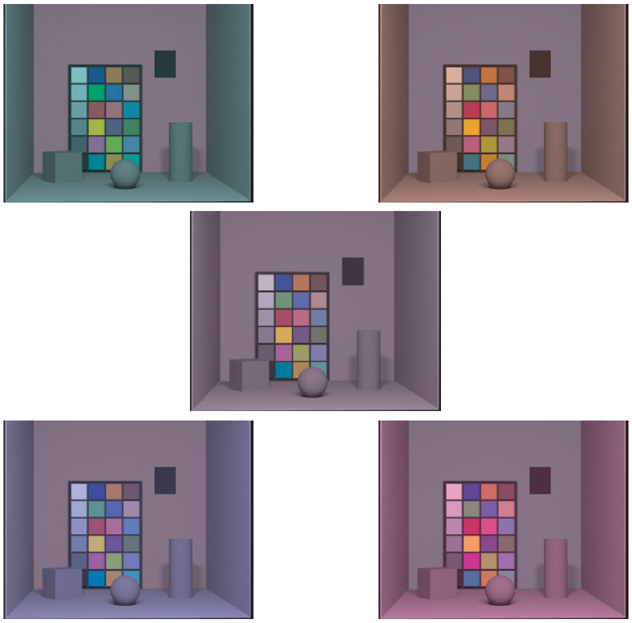

Figure 8. The experimental images used in Experiment 2.

image for the Neutral illuminant was identical to that used in Experiment 1, and because the same observers were used in Experiment 2, data for this image were not collected again.

The achromatic chromaticities (top panel) and equivalent illuminants (bottom panel) for Experiment 2 are shown in Figure 9, and the constancy indices are shown in Figure 10.

The invalid-cue conditions used in Experiment 2 greatly lower the degree of constancy shown by observers. The overall mean constancy index in Experiment 2 was 0.22 (range, 0.10–0.32), compared to a mean index of 0.73 (range, 0.67–0.81) obtained in Experiment 1. This reduction is consistent with the notion that the local surround plays a large role in color constancy. In agreement with the findings of Kraft and Brainard (1999; also Kraft et al., 2002), constancy does not drop to zero. The visual system is able to use cues other than the local surround of the test patch to adjust to the illumination changes across our experimental images.

The range of constancy indices obtained in Experiment 2 was greater than that obtained in Experiment 1. Indeed, a one-way within-observer ANOVA indicated that the *CI*s for Experiment 2 differed significantly with the direction of illumination change ($F(3, 18) = 3.58$, $p < .05$). The ordering of constancy indices for Experiment 2 was Blue_60 > Green_60 > Yellow_60 > Red_60. This is not the ordering we would have expected based on our qualitative analysis of daylight chromaticities.

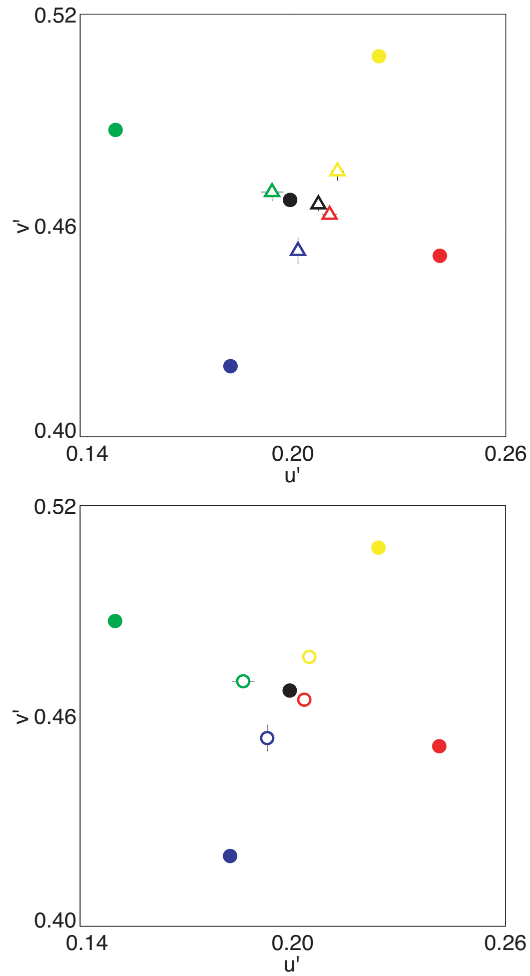

Figure 9. Data from Experiment 2. Same format as Figure 6 above.

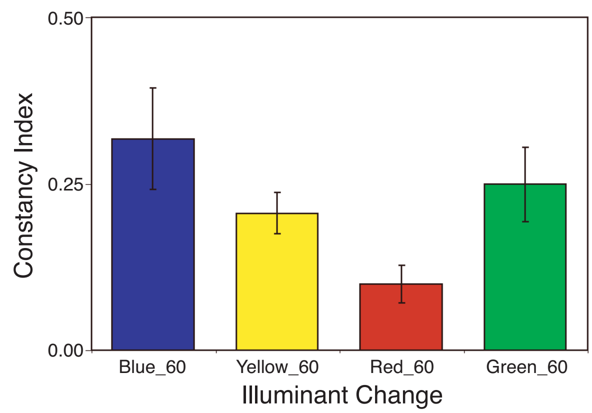

Figure 10. Constancy indices for Experiment 2. Error bars show +/– 1 SEM.

Experiment 3

The results of Experiments 1 and 2 reveal that there are differences in the degree of constancy exhibited for illuminant changes in different color directions. In Experiment 2 these differences were statistically significant. In Experiment 1 the differences were not significant, but the same trend as in Experiment 2 was observed. Because the data were obtained for only one magnitude of illuminant change for each color direction (60 ΔE^* units), it is not possible from the data of Experiments 1 and 2 to determine whether the differences across color direction are intrinsic to the color direction of the illuminant changes or whether they arise because the perceptual magnitude across the different illuminant changes is not precisely equated. That is, if the degree of color constancy depends not only on the direction of illuminant change but also on its magnitude, we might interpret the differences in degree of color constancy as indicating that the four illuminant changes studied had different perceptual magnitudes. Although the illuminant changes were constructed to have equal magnitude in a perceptually uniform color space, the uniformity of this space is at best approximate. This is particularly true for evaluating the size of illumination changes, because the ΔE^* metric is based on judgments of color difference between test patches viewed in surface mode.

We repeated Experiments 1 and 2 with four additional illuminant changes. These shared color direction with the illuminant changes studied in Experiments 1 and 2, but had half the magnitude as measured by the CIELUV E* metric (30 E* units rather than 60). The coordinates of these illuminants are shown in Table 1. Constancy for the four illuminant changes was again assessed with respect to the Neutral illuminant. The light emitted from the monitor from a region on the back wall adjacent to the test patch position was measured after each experimental run. The mean chromaticities and luminances of these measurements are shown in Table 3. Note that the luminance value for the Neutral illuminant is slightly different from the value shown in Table 2 because it is the mean of a subset of the measurements made in Experiments 1 and 2, as only a subset of the original observers participated in Experiment 3.

Illuminant	Back wall	CIE u'	CIE v'	Lum (cd/m^2)
'Neutral'	Normal	0.211	0.470	5.02
'Blue_30'	Normal	0.203	0.448	4.95
'Yellow_30'	Normal	0.226	0.491	5.08
'Red_30'	Normal	0.236	0.464	5.02
'Green_30'	Normal	0.183	0.481	4.73
'Blue_30'	Equated	0.211	0.470	5.03
'Yellow_30'	Equated	0.212	0.472	4.92
'Red_30'	Equated	0.211	0.471	5.05
'Green_30'	Equated	0.205	0.472	4.80

Table 3. The results of measurements of the back wall region of the experimental images are provided for all images used in Experiment 3.

Four of the observers from Experiments 1 and 2 participated in Experiment 3. All were naïve as to the purpose of the experiment except for author PBD. The methods were the same as in Experiments 1 and 2.

The mean achromatic and equivalent illuminant settings are shown for both the valid-cue and invalid-cue conditions in Figure 11. The *CI*s for Experiments 1-3 are shown in Figure 12. The differences in constancy indices

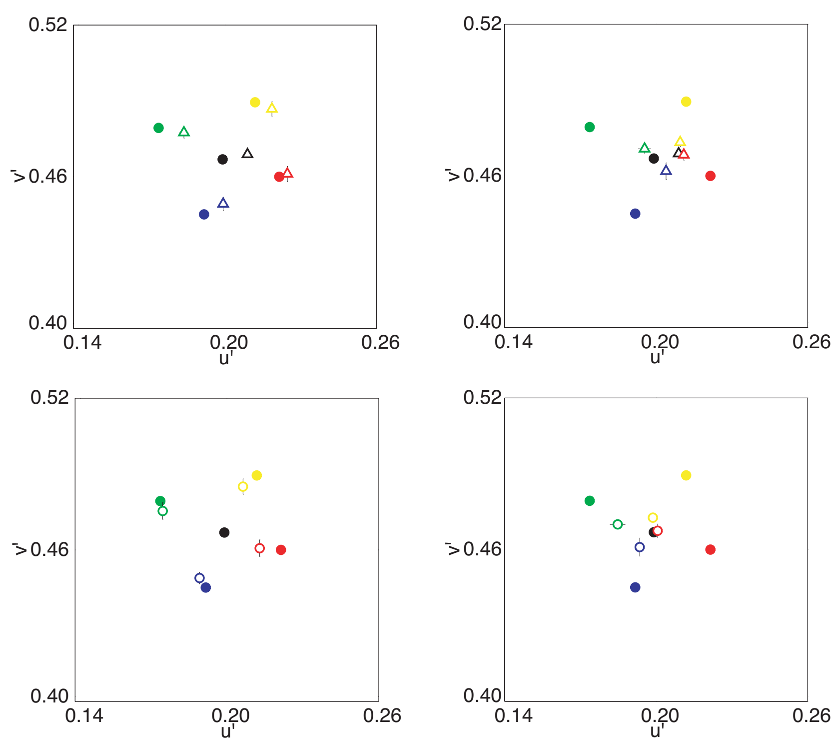

Figure 11. Data from Experiment 3 for valid-cue conditions (left panels) and invalid-cue conditions (right panels). Top panels: Achromatic chromaticities averaged over four observers (open triangles) and chromaticities of corresponding experimental illuminants (solid circles). Bottom panel: Equivalent illuminants derived from the achromatic chromaticities (open circles) and chromaticities of corresponding experimental illuminants (solid circles). Where visible, error bars show +/– 1 SEM.

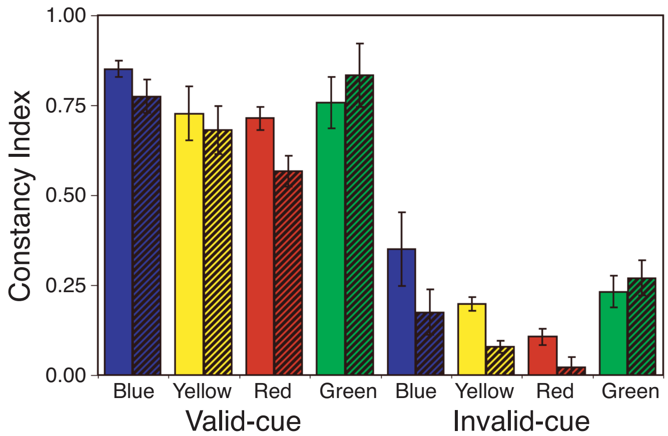

Figure 12. The mean constancy indices obtained in Experiment 3 are shown. The plain bars are the indices for the 60 ΔE * illuminant changes (replotted from Experiments 1 and 2), and the patterned bars are for 30 ΔE * (Experiment 3). The error bars show +/– 1 SEM.

with illuminant change magnitude are generally small, and the pattern of results is similar for both magnitudes, with the exception of a reversal of the magnitude of Blue and Green constancy indices with the change in magnitude. As in Experiments 1 and 2, one-way within-observer ANOVAs, run separately for the valid- and invalid-cue conditions, indicated that the effect of illuminant direction reached statistical significance at the 0.05 level for the invalid-cue condition ($F(3, 9) = 3.87$, $p < .05$) but not for the valid-cue condition ($F(3, 9) = 2.28$, $p = .15$).

To examine the effect of illuminant magnitude, we ran two-way within-observer ANOVAs for data combined from Experiments 1-3, using data from the four observers who participated in all three experiments. The results of this ANOVA are provided in Table 4. The effect of illuminant change magnitude on the *CI*s was not significant for either valid- or invalid-cue conditions. In these ANOVAs, *p* values for the effect of illuminant change direction drop relative to the one-way ANOVAs, so that for invalid-cue conditions, statistical significance is obtained only at the 0.10 level rather than at the 0.05 level. The ANOVA also indicates that the interaction between illuminant change magnitude and direction was not significant.

Supplemental experiments and analyses

In addition to the main experiments, a number of supplemental experiments were conducted. These experiments examined the reliability of observer settings over time and consistency across observers, the effect of changing the basis functions used to generate illuminant spectral power distributions from chromaticity coordinates and luminance, the effect of varying the instructions given to the observers, and the effect of viewing the experimental images monocularly rather than stereoscopically. The data indicate that observer achromatic settings are stable over time, that choice of illuminant basis functions has little effect on the data, that instructions have a small but measurable effect but that this effect does not interact with the effect of the color direction of the illuminant change, and that viewing images monocularly does not affect the achromatic settings. There are systematic individual observer differences. The interested reader is referred to the Appendix B, where each of these experiments is presented in detail. Appendix C presents additional analyses of the data.

2- WAY ANOVA					
Source of Variation	SS	df	MS	F	P-value
VALID-CUE					
Magnitude	0.019	1	0.019	0.637	0.483
Error	0.088	3	0.029		
Direction	0.157	3	0.052	1.871	0.205
Error	0.252	9	0.028		
Interaction	0.052	3	0.017	2.027	0.181
Error	0.077	9	0.009		
INVALID-CUE					
Magnitude	0.058	1	0.058	7.704	0.069
Error	0.023	3	0.008		
Direction	0.215	3	0.072	3.300	0.072
Error	0.196	9	0.022		
Interaction	0.049	3	0.016	1.189	0.368
Error	0.123	9	0.014		

Table 4. Within-observer two-way ANOVAs for Experiments 1-3, for data from the four observers who participated in all experiments. The two factors were the magnitude of the illuminant change (30 and 60 ΔE*) and the direction of the illuminant change (Blue, Yellow, Red, and Green).

Discussion

Empirical summary

The primary purpose of the experiments reported here was to assess whether the visual system's adjustment to changes in illuminant (relative to a Neutral illuminant) depends on the color direction of the illuminant change. This question is of interest because an analysis of the distribution of natural daylights indicates that some illuminant changes are much more likely to occur than others. If the visual system takes advantage of this prior information, intuition suggests that there would be an anisotropy in the degree of color constancy obtained across illuminant directions.

In the main experiments (Experiments 1-3), statistically significant differences in constancy with illuminant change color direction were found for the invalid-cue conditions, but not for the valid-cue conditions. The ordering of constancy across the four illuminant change directions was fairly consistent across all experiments with the most constancy shown for Blue and Green illuminant changes, least for Red illuminant changes, and an intermediate degree for Yellow illuminant changes. The relative degree of constancy for Blue and Green changes varied between experiments.

Supplemental experiments reported in Appendix B provide additional measurements of constancy for a subset of color directions. The ordering of constancy indices found in the supplemental experiments was consistent with the ordering found in the main experiments, although many of these were conducted only for Blue and Red illuminant changes. Supplementary analyses are presented in Appendix C. These also lead to results consistent with the main experiments, with the single exception that a separate analysis of the decremental stimuli for Experiment 1 produced a *CI* for the Yellow illuminant change (0.76) that was slightly lower than that for the Red illuminant change (0.77).

Overall, the most salient effect in our reading of the data is that constancy across the Red illuminant change is less than that for the other directions. The good constancy we measured for Green illuminant changes, particularly relative to that we measured for Yellow changes, does not lend support to the idea that constancy for changes consistent with natural daylight (our Blue and Yellow changes) is better than that for changes inconsistent with natural daylight (our Green and Red changes).

Related studies

Other authors have studied color constancy for different illumination changes. All employed designs analogous to our valid-cue conditions. Indeed, using methods similar to those we employed, but with stimuli consisting of real illuminated objects, Brainard (1998) found no advantage for illumination changes along the blackbody locus compared to those off it. Brainard's (1998) experimental power was reduced by the fact that his study employed only two observers.

Somewhat better constancy for a Blue illuminant change than for a Yellow illuminant change was found by Lucassen and Walraven (1996). They used an index similar to ours to quantify color constancy. The mean *CI* for three observers for their Blue illuminant change was 0.74. For their Yellow illuminant change, it was 0.64. Their results are consistent with ours.

Ruttiger et al. (2001) tested color constancy across illumination changes along the L-M, S, and daylight axes. Although their primary concern was to compare the performance of color normals and color deficient observers, their data showed that the color normals were on average less color constant along the daylight axes. Foster et al. (2003) report similar results and suggest that the lower constancy for daylight changes might be due to reduced inputs from the S-cones to the constancy mechanisms. Both of these studies employed computer-based Mondrian-type stimuli.

Overall, the current literature does not support the idea that the statistics of natural daylight are reflected in the degree of human color constancy with respect to the direction of the illuminant change.

Valid- and invalid-cue conditions

An important feature of our design is that we studied constancy using both valid- and invalid-cue conditions. Kraft and Brainard (1999; Kraft et al., 2002) emphasized the increased empirical power provided by studying invalid-cue stimuli. For example, Kraft et al. (2002) were able to show an effect of scene complexity on color constancy, but this effect was only revealed in their invalid-cue condition. Similarly, we generally find statistically significant effects of illuminant change direction in our invalid-cue conditions. Our use of invalid-cue conditions is closely related to the cue-conflict approach employed by Yang and Maloney (2001) (see Brainard, 2003).

Link to the statistics of daylight

As noted above, our results do not seem compatible with the notion that the likelihood of an illuminant change predicts how color constant the visual system will be for that change. In the measurements of natural daylight spectra, the likelihood of our Blue and Yellow illuminant changes is vastly greater than that of our Red and Green illuminant changes. Consistent with this, constancy across changes in the Red direction is reduced relative to the other three directions. On the other hand, constancy across changes in the Green direction is not similarly reduced. Indeed, in our data it consistently exceeds constancy in the Yellow direction and is comparable with constancy in the Blue direction.

It seems worth considering ways in which our intuition about the importance of prior information might be reconciled with our data. One possibility is that measurements of daylight are not the appropriate database from which to infer the statistics of illuminant changes with which our visual systems must cope. In some scenes, the illumination reflected from an object does not reach that object directly from a light source but instead is reflected indirectly from other objects in the scene. Where indirect illumination plays an important role, the spectrum of the illumination impinging on the objects in the scene may differ considerably from that of the source. Endler reports that this effect can be quite significant in forested areas where tree canopies overlap, and that "forest shade is greenish to yellow-green" (Endler, 1993, p. 10). If our visual systems have evolved or developed in the presence of considerable indirect illumination from foliage, good constancy for Green illuminant changes is less surprising. This observation might explain the asymmetry between constancy across changes in the Green and Red directions. Another possibility along these lines is that exposure to artificial light, whose statistics in our day-to-day environment are not currently well characterized, plays an important role in shaping color constancy. At present, however, these ideas must be taken as speculative, with fuller evaluation awaiting richer measurements of the statistics of the illumination we encounter.

On the other hand, in Appendix B we show that for invalid-cue conditions there is a significant interaction between the direction of the illuminant change and the degree of constancy shown by individual observers. Such an interaction is difficult to explain under the hypothesis that color constancy across illuminant directions is determined by the occurrence statistics of illuminants in natural viewing, because such statistics would seem to be common across observers. This interaction is not, however, something that the current data set allows us to study in detail. To reconcile systematic observer differences with the broad hypothesis that natural image statistics drive the degree of color constancy across illuminant directions would require a theory of how these statistics differ for different observers.

Another possibility that must be considered is that our intuitions about the link between illuminant probabilities and predicted degree of constancy are in error. To quantify these intuitions requires implementation of a Bayesian calculation (e.g., Brainard & Freeman, 1997) that takes into account not only prior probabilities of illuminants but also the prior distribution of surface reflectance functions and the perceptual cost of various constancy failures, followed by an exploration of the effect of varying the prior on predicted performance. Such an exercise is beyond the scope of the present study, but we plan to pursue it in future work.

A final possibility is that our experimental stimuli were not sufficiently natural as to evoke the same performance that the visual system exhibits for natural viewing. We discuss this possibility next.

Use of graphics simulations

To study visual performance as it applies to natural viewing, the experimentalist faces a dilemma. To ensure that the results obtained generalize to situations outside the laboratory, it is desirable to employ stimuli that approximate the richness of natural scenes. To allow accurate specification and manipulation of the stimulus, however, it is necessary to simplify and employ stimuli that capture some but not all aspects of natural viewing.

Many studies of color and lightness constancy employ simulations of rather abstract scenes, flat matte objects viewed under spatially uniform illumination or simple illumination gradients (e.g., Burnham, Evans, & Newhall, 1957; Arend & Reeves, 1986; Brainard & Wandell, 1992; Bauml, 1994). These are simple enough to allow both complete stimulus specification and parametric stimulus manipulation. On the other hand, these stimuli do not look much like natural scenes.

Other studies have employed richer stimuli, consisting of real illuminated objects (e.g., Hochberg & Beck, 1954; Gilchrist, 1977; Brainard et al., 1997; Brainard, 1998; Bloj, Kersten, & Hurlbert, 1999; Rutherford & Brainard, 2002). Results obtained with these stimuli seem more likely to apply to natural viewing. This generalizability is accompanied by less complete stimulus specification and an increase in the difficulty of instrumenting experimental manipulations.

The stimuli used in the present study represent an interesting middle ground. The stimuli consist of digital images displayed on well-calibrated computer-controlled monitors. For this reason, it is straightforward for us to provide a complete specification of what the observers saw. Because the simulated scenes are specified in software, manipulation of the stimuli is more easily accomplished than when one experiments with physical illuminated surfaces. At the same time, the physics-based rendering software used allows our simulated scenes to appear similar to photographic images of actual scenes.

Our experimental procedures and data analysis are similar to those employed by Kraft and Brainard (1999; Kraft et al., 2002), except that their stimuli consisted of real illuminated objects. The levels of constancy exhibited in our experiments are similar to those they found. For valid-cue conditions, Kraft and Brainard (1999) found a mean constancy index of 0.83 (average of 4 observers), whereas Kraft, Maloney, and Brainard (2002) found a mean constancy index of 0.85 (average of 10 observers). This compares to our average index of 0.72 from Experiment 1 and the valid-cue condition of Experiment 3. For the invalid-cue conditions, Kraft and Brainard found a mean constancy index of 0.53 (mean of 4 observers), whereas Kraft et al. (2002) found a mean index of 0.25 (average of 10 observers). This compares to our average index of 0.19 from Experiments 2 and the invalid-cue conditions of Experiment 3. Given that there are a number of differences in the details between the various studies (e.g., illuminants, surface reflectance functions of objects in the scene, size of the scenes, and identity of observers), we feel that the overall similarity in the constancy indices across the experiments with real and simulated images provides some assurance that the simulated images we used provide a reasonable laboratory model for natural viewing. A more definitive statement awaits direct empirical comparisons between performance measured for real scenes and for simulations of these scenes.

Apart from the issue of simulation, it is worth noting that the simulated illuminant intensities were much lower than those of many daylights. In spring 1999, one of the authors (PBD) made some daylight measurements at the University of California, Santa Barbara. In shadow, the luminance of the light reflected from a perfect diffuser was 1650 cd/m^2, and in direct sunlight the corresponding luminance was 27500 cd/m^2. This compares to the luminance level of 25 cd/m^2 simulated in the present experiments. This low level was used because of limits on the light output of our CRT displays. Not much is known about how color constancy is affected by the overall level of the light. In one recent study, however, Delahunt and Brainard (2000) found similar results for an asymmetric matching task performed using a stimulus presented on a CRT (mean luminance level of 14 cd/m^2) and on a rear-projection system (mean luminance level of 590 cd/m^2). As with the issue of how stimulus complexity affects constancy, firmer conclusions about the effect of overall light level will require additional experimentation.

Finally, it should be noted that our study employed essentially a single spatial arrangement of objects in the simulated scene. An interesting open line of experimentation is to understand how different choices of scene objects affect color constancy. Such studies are enabled by the use of synthetically produced stimuli, because with such stimuli, variation in scene composition becomes practical on a trial-by-trial basis.

Acknowledgments

We thank Jeff DiCarlo and Brian Wandell for making their dataset of daylight measurements available to us. Jerry Tietz and Philippe Longère provided technical advice and support. This research was supported by National Institutes of Health Grant EY 10016.

Commercial relationships: none.

Corresponding author: Peter B. Delahunt.

Email: pbdelahunt@ucdavis.edu.

Appendix A: Methods details

This appendix complements the Methods section in the body of the text by providing a number of additional methodological details.

RADIANCE

RADIANCE produces images from a text-based *scene description*. The foundation of the description is the *rendering space* that specifies the dimensions of a three-dimensional volume that contains the simulated light sources, objects, and observer. Within the rendering space, light sources are specified by their position, size, and power in three color bands. The color bands are referred to as red, green, and blue (RGB). Similarly, objects are specified by their position, shape, size, and reflectance properties. In the present experiments, all objects were defined to be Lambertian reflectors. Given this, the reflectance parameter in the scene description that can vary from object to object is the percentage of incident light reflected in the red, green, and blue color bands. Below (see Spectral rendering) we describe how we extended this RGB model to provide more accurate rendering of color.

To compute an image from the scene description, it is necessary to specify where the scene is viewed from. Once the position and direction of view are provided, RADIANCE uses the scene description to compute RGB light intensities at each image location. To generate left and right eye images from the scene description, we used two different viewpoints separated horizontally by 6 cm (2.4 in.), roughly the distance between adult observers' eyes.[4]

The ray-tracing algorithms implemented in RADIANCE simulate the propagation of light from source to object to image. For real scenes, some rays reflect off multiple objects before they reach the observer. Tracing rays through arbitrary multiple reflections is computationally intractable, and RADIANCE provides a parameter that determines the maximum number of reflections to simulate during the rendering process. For our images, this parameter (called INDIRECT in the RADIANCE documentation) was set to 2.

The RADIANCE software requires that an exposure parameter is set before creating an image. This parameter determines the relation between the overall intensity of the simulated illuminant and the overall intensity of the rendered image. To set this parameter so that the rendered scenes corresponded to a known physical illuminant intensity, images were rendered for a range of exposure settings, and a Photo Research PR-650 spectraradiometer was used to measure a reference location in each image. Because the simulated reflectance of the reference location was known, this allowed us to choose an exposure parameter that led to the desired simulated illuminant intensity.

The RADIANCE resolution parameter was set to 1092 and the output images had a resolution of 805 (h) × 1092 (w). The images were presented at this resolution setting on the experiment monitors. The RADIANCE scene description files, illuminant and object spectra, and rendering parameters can be viewed by clicking here (supplementary material).

Spectral rendering

To ensure colorimetric accuracy of our rendered images, we employed a spectral rendering procedure. The spectra for all illuminants and surfaces were stored in a single supplementary text file. A master scene description was created, where each illuminant and surface were given dummy RGB values. A PERL script was used to create 31 scene descriptions from the master. For each scene description, the dummy RGB values were replaced with actual spectral values taken from the supplementary text file. In these scene descriptions, RGB values were set equal to each other (i.e., R = G = B) for all illuminants and surfaces. The PERL script then ran RADIANCE 31 times, once for each scene description. This procedure provided us with 31 simulated monochromatic images at wavelengths between 400 nm and 700 nm in 10 nm steps.

Stereo apparatus

The stereo apparatus is illustrated in Figure 13. Two 21 in. display monitors (Hewlett Packard Model P1110) were each driven by separate graphics cards (Radius, 10-bit DACs), both controlled by a single host computer (Apple PowerMac G3). The monitors were placed at an optical distance of 36 in. from the eyes of the observer and placed so that the center of the screens was at the observer's eye level. The monitors were oriented such that the display surface was perpendicular to the optical axis. The beamsplitters had a transmission efficiency of 49% and a reflection efficiency of 39%. The angle of the beamsplitter with respect to the light source had a negligible effect on these efficiencies. Beamsplitters were used rather than mirrors to facilitate alignment of the apparatus. All calibrations were performed in situ, so that the spectral reflectance function of the beamsplitters was accounted for.

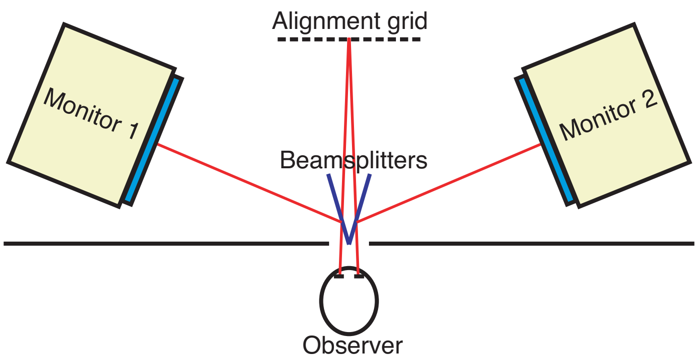

Figure 13. Schematic of the stereo viewing apparatus.

The apparatus was placed in a dark room, and room surfaces visible from the observer's vantage point were covered with either black matte paper or black cloth. The observer viewed the monitors through two square apertures measuring 1.75 in. square. A chin rest was used to stabilize the observer's head position. The experiments were conducted with the room lights off.

The beamsplitters were placed so that the reflected images from each monitor converged at a point in the observer's frontal plane at an optical distance of 36 in. from the eyes. The individual beamsplitters were oriented as close as possible to vertical with any errors in this adjustment being compensated for by appropriately positioning the monitors. The angle between the beamsplitters was 62.5 deg, which places the monitors forward of the observer. This allowed us to build an occluding wall that prevented the observers from seeing the apparatus.

Once the basic geometry of the apparatus was established, fine adjustments to the alignment of the two monitors were made under visual guidance. A cardboard alignment grid measuring 17 in. (w) × 13 in. (h) was placed directly in front of the viewing position at a distance of 36 in. The grid was located in the observer's mid-sagittal plane at eye level and contained horizontal and vertical lines spaced 1 in. apart. The same grid was simulated on both monitors and the alignment was refined by adjusting the size, tilt, and distortion of the simulated grid on each screen, using built-in monitor controls. Adjustments were made until the simulated grids in each eye were superimposed over the alignment grid. During the experiments, the alignment grid was obscured from view using a matte black occluder. The alignment of the apparatus was checked every two weeks.

To verify that our apparatus produced realistic renditions of scene dimensions, a cardboard box (side 8 in.) with one open face was constructed and positioned in front of the observer so that the back wall was 36 in. from the observer's eyes. This box was then simulated using the RADIANCE software, and the resulting stereo pair displayed on the apparatus. Two observers verified informally that the perceived distances to the various areas of the rendered box were in general agreement with the perceived distances to the corresponding areas of the real box.

Scene dimensions and content

The visual angle of objects in the scene may be computed from their simulated sizes and simulated distance from the observer. Within the rendering space, a box was created measuring 10 in. (w) × 14 in. (h) × 10.5 in. (l). This box was placed at a height of 6 in. from the bottom of the rendering space. It was positioned directly in front of the viewer at the furthest point from the viewer in the rendering space. The scene was viewed at a height of 10 in. from the bottom of the rendering space. The left and right scenes were rendered from two different viewpoints separated horizontally by 2.4 in.[4] Not all of the RADIANCE scene was visible when rendered on the monitors (see Figure 3). When rendered, the visible portion of the box subtended 21.9 deg horizontally and 17.8 deg vertically.

The simulated light source was of size 10 in. (w) × 1 in. (h) × 2 in. (l) and was positioned in a horizontally central position 26 in. from the back wall of the room at a height of 18 in. There were four simulated objects in the scene. The cube was of side 1.5 in., was rotated 30 deg counterclockwise, and placed on the left side of the room 5 in. from the back wall. The sphere was of radius 0.7 in. and placed in a central position 6 in. from the back wall. The dimensions of the cylinder were 3 in. (h) with radius 0.6 in., and it was placed on the right side of the room at a distance 4 in. from the back wall. The fourth object was a simulated Macbeth color checker chart (GretagMacbeth LLC) with each square being of side 1 in. with a .125 in. trim. The test patch on the back wall was 1.6 in. (v) × 1.25 in. (h) and placed at a height of 6 in. above the room floor and 1.5 in. to the right of the horizontal center of the back wall.

The full spectra of the simulated illuminants are provided as part of the supplementary material. The spectra of the Neutral, Blue_60, Blue_30, Green_60, Green_30, Yellow_60, Yellow_30, Red_60, and Red_30 spectra were constructed from their chromaticities and luminances by constraining them to lie within the three-dimensional linear model for daylights defined by the CIE (1986). When we applied this procedure to the chromaticity and luminance of the Green_60 and Green_30 illuminants, however, we found that the constructed spectrum had negative power at some wavelengths, a property we did not view as desirable. Therefore, the spectrum of these illuminants was constructed by constraining it to lie within a linear model defined by the red, green, and blue phosphor emission spectra of a monitor measured in our laboratory. A control experiment described in Appendix B investigates the effect of varying the linear model used to construct full spectra from illuminant chromaticity and luminance.

Monitor calibration

The output of the rendering process is the L-, M-, and S-cone excitation coordinates desired at each image location. Conversion between these cone coordinates and monitor settings was achieved using the general model of monitor performance and calibration procedures described by Brainard (1989). Monitor calibrations were performed every two weeks using the spectraradiometer. Spectral measurements were made at 4-nm increments between 380 and 780 nm but interpolated with a cubic spline to the CIE-recommended wavelength sampling of 5-nm increments between 380 and 780 nm. CIE XYZ and chromaticity coordinates were computed with respect to the CIE 1931 color-matching functions. The spectral power distribution of each phosphor was measured at a range of intensity levels to measure and correct for the nonlinear relation between digital input and light intensity output characteristic of CRT monitors.

To correct the data for any small violations of the calibration assumptions, the observer's settings were replayed after each session and measured directly using the radiometer. This provided direct measurements of the achromatic adjustments. The simulated illuminant was also assessed by measuring the light emitted from an area just below the test patch and then converting this measurement using the known simulated surface reflectance function of that location. To speed up this process, only one monitor was measured after each experimental run. Inconsistencies between the two monitors were small, with xy chromaticity differences generally less than .004 and luminance differences less then 5%. We believe the luminance differences arose because of differences in the spatial inhomogeneities of the two monitors. To correct the measurements made on one monitor for the luminance differences between the two monitors, a correction factor was calculated for each measured area and used to estimate the mean luminance of the two monitors. The correction factors were re-measured after each monitor calibration.

Appendix B: Supplemental experiments

Individual observers

To verify the stability of observers' achromatic settings over time, we re-measured achromatic chromaticities for the Neutral, Blue_60, and Red_60 illuminants for both valid and invalid conditions. These measurements were made about 2 months after the data reported above for Experiments 1 and 2 were collected. The same seven observers used in Experiments 1 and 2 participated in this replication. Figure 14 shows the mean equivalent illuminant settings for the original and repeated experiment for both the valid-cue and invalid-cue conditions. Figure 15 shows the mean *CI*s.

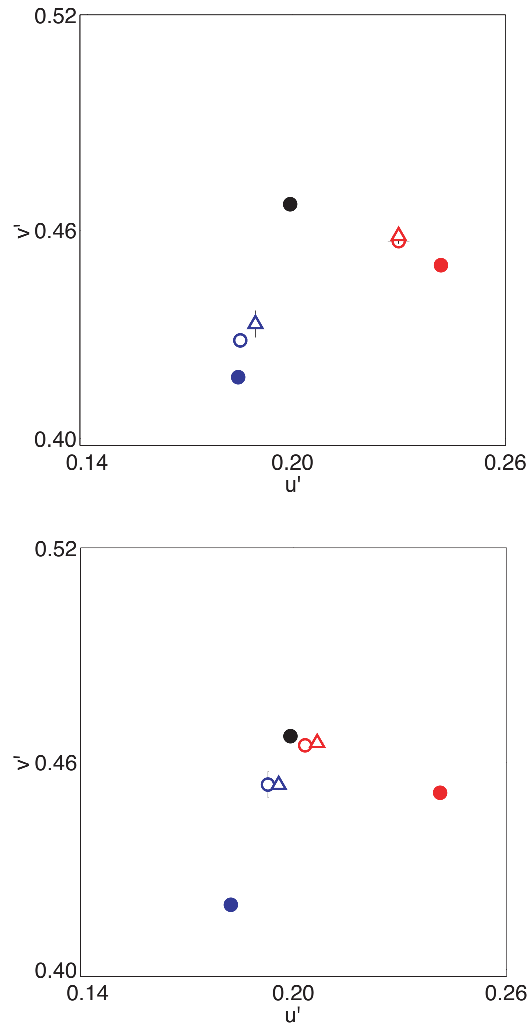

Figure 14. The top panel shows the valid-cue conditions and the bottom panel shows the invalid-cue conditions in u'v' chromaticity coordinates. The mean settings (open symbols) are plotted against the illuminants (filled circles). The open circles are the original settings and the open triangles are the repeats made about 2 months later. The error bars are +/– 1 SEM.

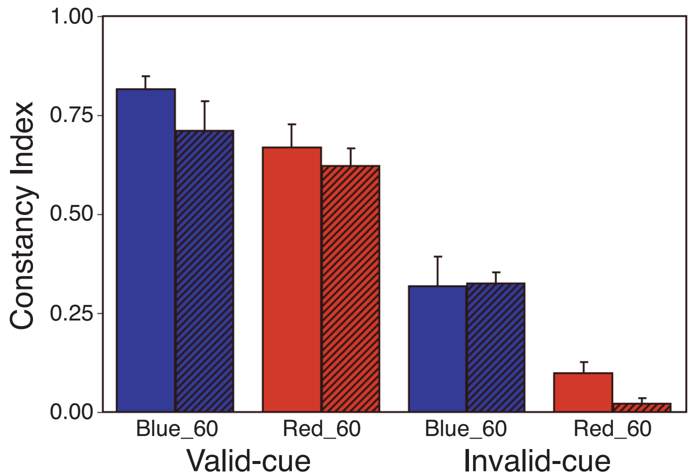

Figure 15. The plain bars are the constancy indices from the original experiment and the patterned bars are the replication. The error bars are 1 SEM.

Figure 16 plots the individual observer *CI*s for the repeated experiments against the *CI*s from the original experiment. The diagonal line indicates equality and the plotted points fall close to this line indicating that individual observer settings are stable over time. The correlation between the original and repeated measurements is 0.90. These data, together with the similarity in the mean results, indicate that our measurements are quite reliable.

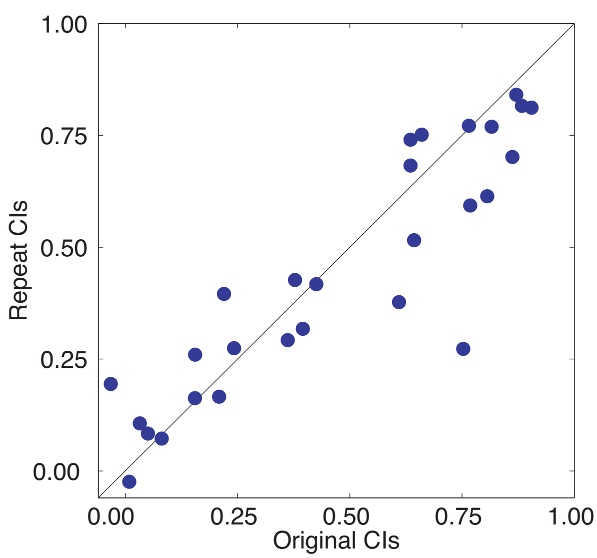

Figure 16. The constancy index measurements from the repeated experiment are plotted against the original experiment. The diagonal line indicates equality and with a few exceptions the plotted points fall close to this line. The correlation across measurements was 0.90.

2- WAY ANOVA

Source of Variation	SS	df	MS	F	P-value
VALID-CUE					
Illuminant Change	0.096	1	0.096	7.072	0.019
Error	0.190	14	0.014		
Observer	0.304	6	0.051	3.723	0.020
Error	0.190	14	0.014		
Interaction	0.064	6	0.011	0.789	0.593
Error	0.190	14	0.014		
INVALID-CUE					
Illuminant Change	0.315	1	0.315	52.29	0.000
Error	0.085	14	0.006		
Observer	0.105	6	0.018	2.897	0.047
Error	0.085	14	0.006		
Interaction	0.173	6	0.029	4.785	0.007
Error	0.085	14	0.006		

Table 5. Two-way ANOVA with illuminant change and observer as the factors. Two illuminant change directions (Blue and Red) were used, and data were available for four observers. The replication was repeat measurements for each observer.

Table 5 presents results of a two-way ANOVA with observer and illuminant change direction as the factors. The repeated measurements for each observer were used as the replication. Consistent with our group analysis in the main part of the paper, this ANOVA shows a significant effect of illuminant change (here Blue vs. Red only) for both valid- and invalid-cue conditions. The ANOVA also indicates that there are significant individual observer differences and (in the invalid-cue condition) a significant illuminant change by observer interaction.

A detailed analysis of the individual observer differences is beyond the scope of this work. Indeed, such an analysis would be more straightforward to carry out had we obtained across-session replications for all observers in all conditions. The reader interested in individual observer differences is referred to the supplementary material, where the individual observer data is tabulated.

Instructional effects

The perceptual criteria used by observers can affect the results obtained in color appearance experiments. In asymmetric matching experiments, Arend and Reeves (1986) and Bäuml (1999) used instructions to influence observers' perceptual criteria. They distinguished appearance matches, where the observer was instructed to judge the appearance of the light reaching the eye, from surface matches, where the observer was instructed to judge the identity of the simulated surface. They found that constancy was substantially lower for appearance match instructions than for surface match instructions. The conditions under which such instructional manipulations can have a large effect on appearance data are not well established. We tested the effect of instructions for our experiments.

The methods were the same as for Experiment 1 except for the following. Ten additional naïve observers participated in this experiment. Five were males and five were females ranging in age from 18 to 27 years. All had corrected visual acuity better than 20/20 and were stereo normal. Five were randomly assigned to the appearance condition, and five to the surface condition. These conditions were defined by the instructions given to the observers. Each set of instructions was accompanied by a demonstration. The demonstration consisted of shining a tungsten light with and without a blue filter onto white paper. The relevant part of the instructions and a description of the demonstration used for the two conditions follows.

Appearance condition instructions: "The display simulates objects that are illuminated by light. We want your judgment to be made about the color of the light reflected to your eye, not what color of paper the rectangular patch looks like it's made out of. For example, if the patch looks blue because there seems to be blue light falling on gray/white paper, adjust the patch until the blue sensation is gone."

Appearance condition demonstration: "This is a white piece of paper (show white paper under tungsten light). When I shine a blue light on it you can see that the light reflected to your eyes is blue, but you can also see that the paper is white paper (show white paper under blue light). We want you to adjust the patch until the light reflected to your eye appears achromatic, that is until the color disappears. Don't worry about what color the paper looks."

Surface condition instructions: "The display simulates objects that are illuminated by light. We want you to adjust the test patch until it looks like it is made from gray/white paper. A range of illuminants (lights) will be used in various conditions. Each time, we want you to adjust the patch until it looks like a gray/white piece of paper, no matter what color the illuminant is."

Surface condition demonstration: "What color is this paper? (Show white paper under tungsten light - expected response: "white"). Now what color is the paper now? (Show white paper under blue light - expected response: "Still white"). But here the light reaching your eye is bluish. We want you to base your judgment on the appearance of the paper. It may be that when your adjustment is finished, you perceive some color in the test patch because it appears to be gray or white paper under colored light. This is OK."

Measurements were made for the Neutral, Blue_60, and Red_60 illuminants for both valid- and invalid-cue conditions.

The mean equivalent illuminant settings are shown in Figure 17. The *CI*s are shown in Figure 18. For comparison purposes, the *CI*s from the Blue and Red conditions from

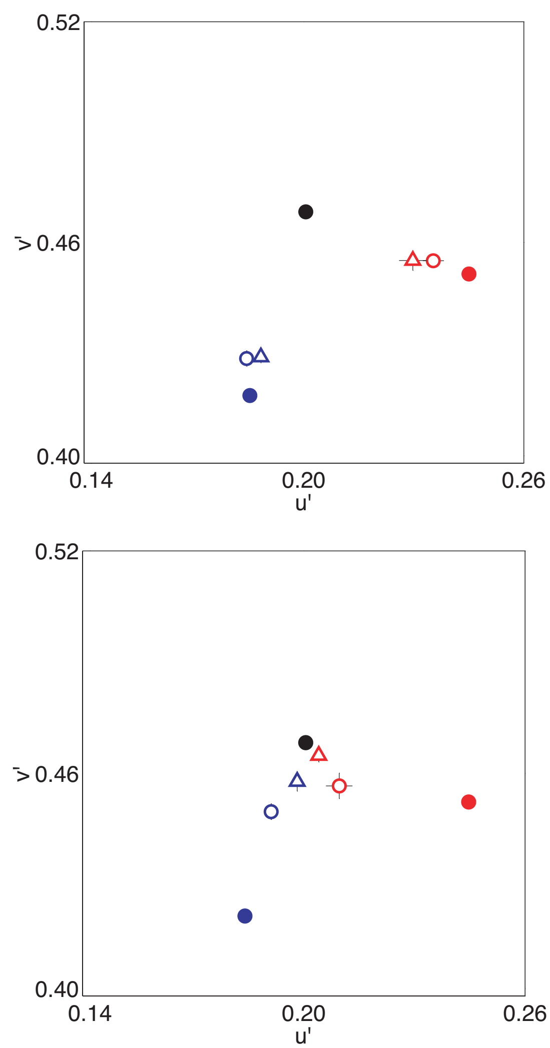

Figure 17. Effect of instruction. The mean settings for the "surface" conditions (open circles) and "appearance" conditions (open triangles) are shown together with the test illuminants (filled circles). The top panel shows the valid-cue conditions and the bottom panel shows the invalid-cue conditions in u'v' chromaticity coordinates. The error bars are +/– 1 SEM.

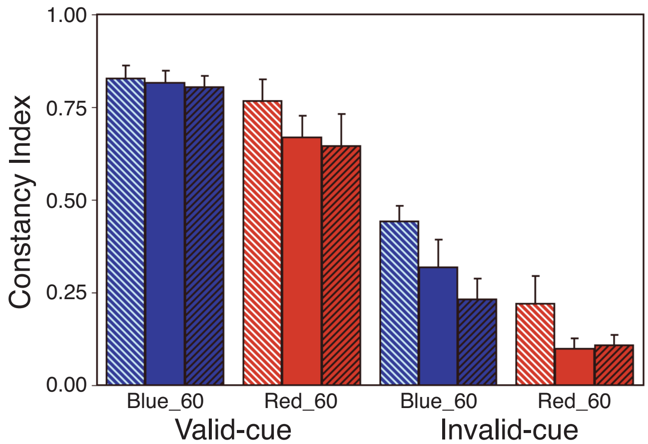

Figure 18. The constancy indices for the "surface" instructions condition (white diagonals) and the "appearance" instructions condition (black diagonals) are compared to the original experiment (plain). The error bars are 1 SEM.

Experiments 1 and 2 are also shown. The differences in *CI*s for the two types of instructions are generally small. The differences are nonsignificant for the valid-cue conditions, and significant for the invalid-cue conditions (see ANOVA, Table 6). The results suggest that the type of instructions used has a small but consistent effect on the levels of constancy. Surface instructions led to higher levels of constancy

2- WAY ANOVA

Source of Variation	SS	df	MS	F	P-value
VALID-CUE					
Instructions	0.026	1	0.026	1.111	0.308
Error	0.371	16	0.023		
Illuminant Change	0.060	1	0.060	2.589	0.127
Error	0.371	16	0.023		
Interaction	0.012	1	0.012	0.518	0.482
Error	0.371	16	0.023		
INVALID-CUE					
Instructions	0.132	1	0.132	6.778	0.019
Error	0.311	16	0.019		
Illuminant Change	0.149	1	0.149	7.670	0.014
Error	0.311	16	0.019		
Interaction	0.012	1	0.012	0.622	0.442
Error	0.311	16	0.019		

Table 6. Two-way between-observers ANOVA for the instructions control experiment. The two factors were the instruction (appearance/surface) and the direction of the illuminant change (Blue, Yellow, Red, and Green).

than appearance instructions. In the main experiments, we used neutral instructions that simply asked the observers to make the test area appear achromatic. The *CI*s from the main experiments tend to fall between the *CI*s for those obtained with the explicit surface and appearance instructions. Although differences were found in the present control experiment, the differences are rather small and do not interact with the direction of the illuminant change. It does not seem likely that the conclusions we draw about the effect of illuminant direction depend on the instructions given to observers.

It is useful to remember that there are substantial differences in design between our experiments and those of previous studies (e.g., Arend & Reeves, 1986; Bauml, 1999) that have demonstrated large instructional effects. Our experiments study color constancy across changes in scenes that occur over time. The previous works studied simultaneous constancy, where scenes rendered under two different illuminants were simultaneously visible. The presence of two explicitly visible illuminants may have allowed subjects in previous experiments access to strategies not as readily available in our experiments, where no simultaneous comparison of illumination was possible.

Stereoscopic viewing

The main experiments were conducted using scenes viewed stereoscopically. A control experiment to study the effect of stereoscopic viewing was conducted. Measurements were made for the Neutral, Blue_60, and Red_60 illuminants when the observers viewed the images monocularly with their right eye only. Five of the original seven observers participated in this experiment. All were naïve as to the purpose of the experiment except for the author PBD.

The data (equivalent illuminants and *CI*s) are shown for the valid-cue and invalid-cue conditions (Figures 19 and 20). A two-way within-observers ANOVA indicated that *CI*s are not statistically different for stereoscopic and monocular viewing (see ANOVA, Table 7). The results of the main experiment do not appear to depend on stereoscopic viewing. Observers did report that the task was much less pleasant to perform monocularly. Note also that one of the observers was stereoblind.

In the ANOVA reported in Table 7, the color direction of the illuminant change (Blue or Red) led to significantly different levels of constancy for the valid-cue condition, and close to significantly different levels for the invalid-cue conditions. The effect of illuminant change did not interact with viewing mode (stereoscopic versus monocular).

Illuminant basis functions

In the main experiments, CIE daylight basis functions were used to construct most of the illuminant spectra. For the green illuminants, however, the desired chromaticities constructed from these basis functions would have had negative power at some wavelengths. Because of this, the Green_60 and Green_30 illuminant spectra used in Ex-

2- WAY ANOVA					
Source of Variation	SS	df	MS	F	P-value
VALID-CUE					
Illuminant Change	0.081	1	0.081	18.795	0.012
Error	0.017	4	0.004		
Viewing Condition	0.026	1	0.026	3.581	0.131
Error	0.029	4	0.007		
Interaction	0.012	1	0.012	0.849	0.409
Error	0.055	4	0.014		
INVALID-CUE					
Illuminant Change	0.346	1	0.346	5.380	0.081
Error	0.257	4	0.064		
Viewing Condition	0.013	1	0.013	1.849	0.245
Error	0.027	4	0.007		
Interaction	0.004	1	0.004	0.287	0.621
Error	0.062	4	0.016		

Table 7. Two-way within-observers ANOVA for the viewing condition control experiment. The two factors were viewing condition (monocular /binocular) and the direction of the illuminant change (Blue, Yellow, Red, and Green).

periments 1-3 were constructed from monitor basis functions, as described in Methods. A control experiment was run to determine whether the basis functions used to construct illuminants of a given chromaticity and luminance influenced the results.

The same methods were used as in Experiments 1 and 2, and the same seven observers participated. The images were synthesized using versions of the Blue_60, Yellow_60, and Red_60 illuminants constructed from the monitor basis functions. The Green_60 illuminant was not included, because this illuminant could not be synthesized in a physically realizable manner with respect to the CIE daylight basis functions.

Figure 21 shows the illuminants and equivalent illuminant settings for both valid-cue and invalid-cue conditions. The *CI*s are shown in Figure 22. A two-way within-observer ANOVA indicated that there are no statistically significant differences in the data obtained with illuminants constructed with the two different sets of basis functions (see ANOVA result, Table 8).

The observant reader will note that the illuminant chromaticities differ slightly between the original images (solid circles) and the images synthesized for this experiment (solid triangles). The illuminants were constructed to have the same chromaticity and luminance as reflected directly from a perfect diffuser in isolation. The illuminant

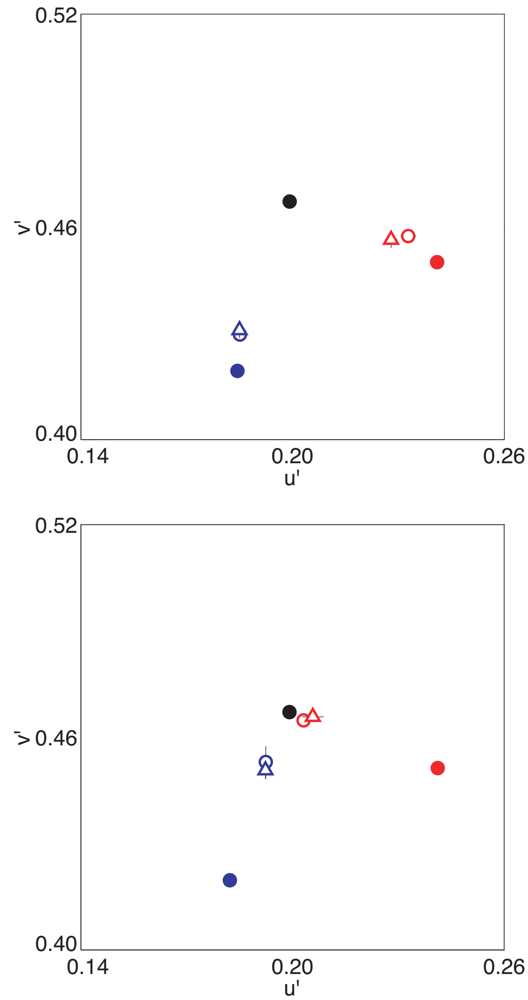

Figure 19. The top panel shows the valid-cue conditions and the bottom panel shows the invalid-cue conditions in u'v' chromaticity coordinates. The mean settings for the binocular viewing (open circles) are compared to the monocular viewing settings (triangles). The illuminants are plotted using filled circles. The error bars are +/– 1 SEM.

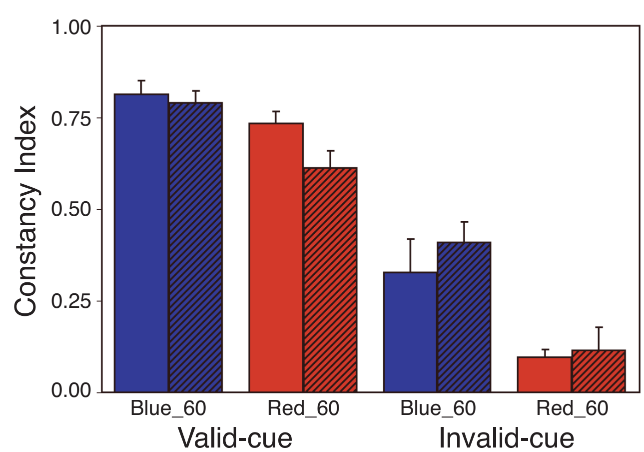

Figure 20. The constancy indices for the binocular conditions (plain bars) are compared to the monocular conditions (patterned bars). The error bars are 1 SEM.

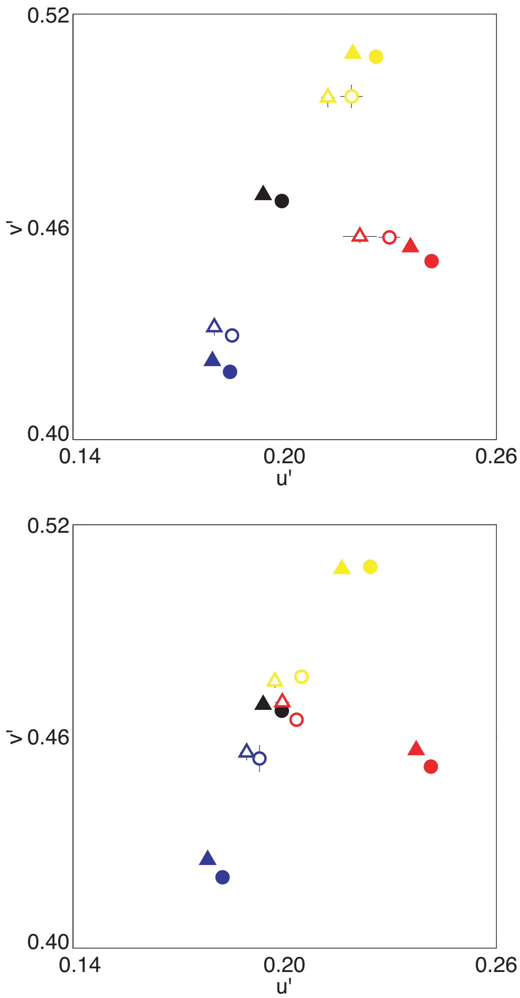

Figure 21. The top panel shows the valid-cue conditions, and the bottom panel shows the invalid-cue conditions in u'v' chromaticity coordinates. The circles show the CIE basis functions condition, and the triangles show the monitor basis functions condition. The error bars are +/– 1 SEM.

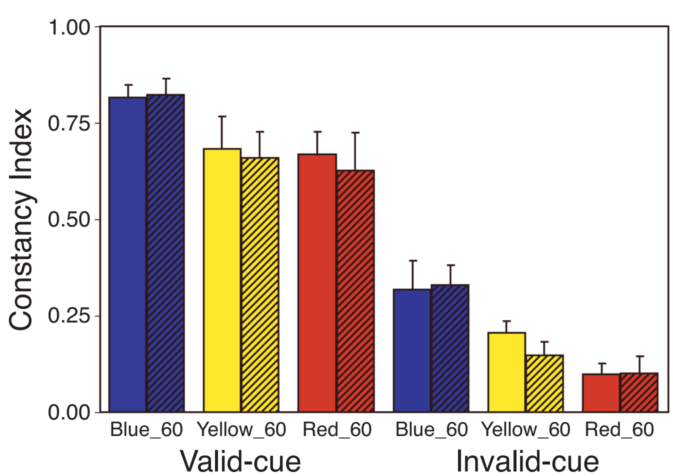

Figure 22. The constancy indices for the CIE basis function condition (plain bars) are compared with the monitor basis functions condition (patterned bars). The error bars are 1 SEM.

2- WAY ANOVA

Source of Variation	SS	df	MS	F	P-value
VALID-CUE					
Basis functions	0.004	1	0.004	2.751	0.148
Error	0.009	6	0.002		
Illuminant Change	0.239	2	0.119	2.662	0.110
Error	0.538	12	0.045		
Interaction	0.005	2	0.002	0.364	0.702
Error	0.074	12	0.006		
INVALID-CUE					
Basis functions	0.003	1	0.003	0.209	0.664
Error	0.073	6	0.012		
Illuminant Change	0.360	2	0.180	8.488	0.005
Error	0.255	12	0.021		
Interaction	0.010	2	0.005	1.253	0.321
Error	0.049	12	0.004		

Table 8. The two-way ANOVA results are shown for the basis function control experiment. The two factors were the type of basis function used (daylight and monitor) and the direction of the illuminant change (Blue, Yellow, Red, and Green).

chromaticities plotted are the ones we measured directly in the experimental images. Presumably the small differences arise because the graphics rendering program synthesizes the flow of light through multiple reflections in the scene. These small shifts in illuminant chromaticities are taken into account via our computation of constancy indices.

In the ANOVA reported in Table 8, the color direction of the illuminant change (Blue, Yellow, or Red) led to significantly different levels of constancy for the valid- and invalid-cue conditions. The effect of illuminant change did not interact with the type of basis function used (daylight vs. monitor).

Appendix C: Supplemental analyses

Decrements versus increments

Previous authors have noted differences in achromatic settings and asymmetric matches depending on whether the test stimuli were increments or decrements (e.g., Mausfeld & Niederee, 1993; Chichilnisky & Wandell, 1996, Chichilnisky & Wandell, 1999; Mausfeld, 1998; Schirillo, 1999a, 1999b; Delahunt & Brainard, 2000; Bauml, 2001). In the current experiments, we used four different test lu-

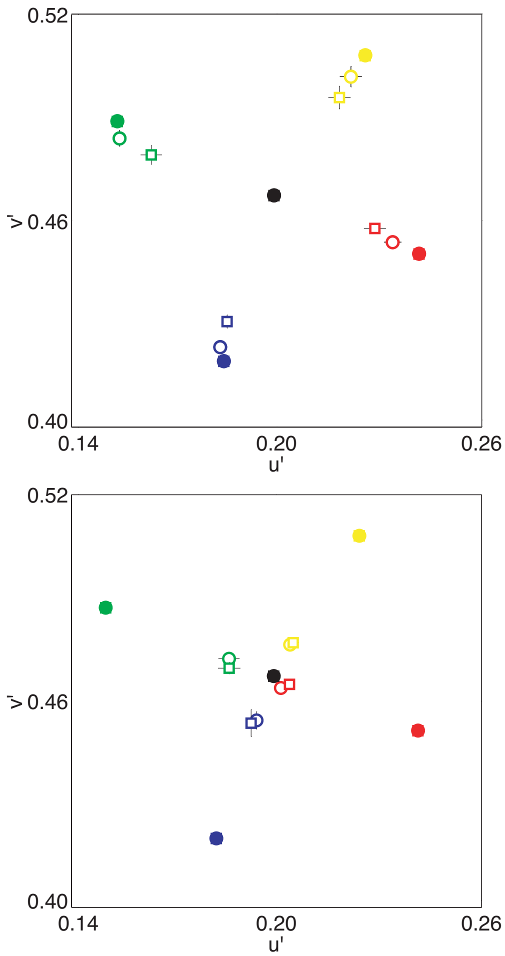

Figure 23. The equivalent settings (open symbols) are shown for both decrements (circles) and increments (squares) for the 60 ΔE* illuminant conditions. The closed circles are the test illuminants. The top panel shows the settings for the valid-cue conditions, and the bottom panel shows the settings for the invalid-cue conditions. The error bars are +/– 1 SEM.

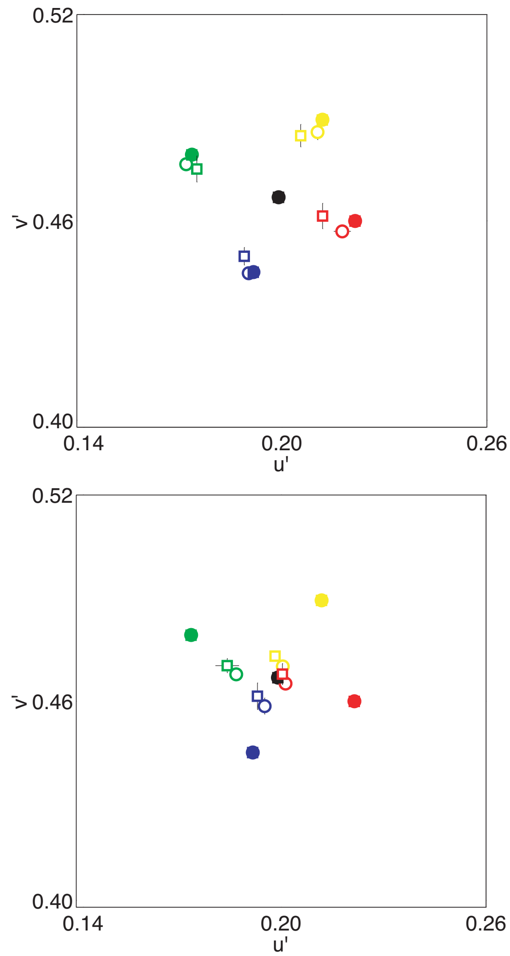

Figure 24. The equivalent settings (open symbols) are shown for both decrements (circles) and increments (squares) for the 30 ΔE* illuminant conditions. The closed circles are the test illuminants. The top panel shows the settings for the valid-cue conditions and the bottom panel shows the settings for the invalid-cue conditions. The error bars are +/– 1 SEM.

minance values, two of which were below the luminance of the local surround, and two of which were above (see Experimental Procedure). Figure 23 shows the equivalent illuminants separately for decrements and increments for both valid- and invalid-cue 60 ΔE^* illuminant conditions. Figure 24 shows the results for the 30 ΔE^* conditions. The CIs are shown in Figure 25 and 26 for the 60 ΔE^* and 30 ΔE^* conditions, respectively. There is a clear difference between increments and decrements for the valid-cue condition, but this difference is not apparent in the invalid-cue conditions. For the valid conditions, decrements lead to better constancy than increments, a result in accord with Bauml (2001). For both valid- and invalid-cue conditions, the general pattern of the effect of illuminant direction is generally the same for increments and decrements.

Constancy index calculation

The constancy index indicates the degree of color constancy across changes in illumination. The calculation is made in two steps:

1. The achromatic settings are recentered so that the settings made under the reference illuminant coincide exactly with the chromaticity of that illuminant. The recentering procedure can be thought of as a prediction of how the achromatic settings data from all illuminant conditions would have looked had the surface that appeared achromatic under the reference illuminant been non-selective. After the recentering is performed on the dataset, the achromatic settings made under the test illuminants are referred to as the *equivalent illuminants*.

2. The constancy index calculation described by Equation 2 above is applied.

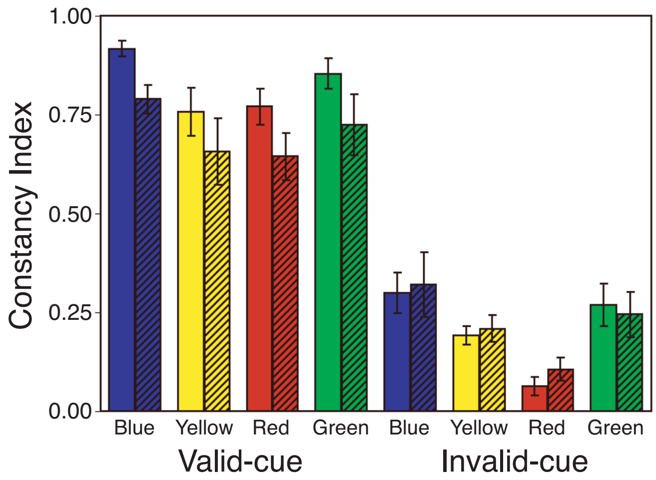

Figure 25. The constancy indices are shown for the decrements (solid bars) and increments (patterned bars) for the 60 ΔE* illuminant conditions. The error bars are +/- 1 SEM.

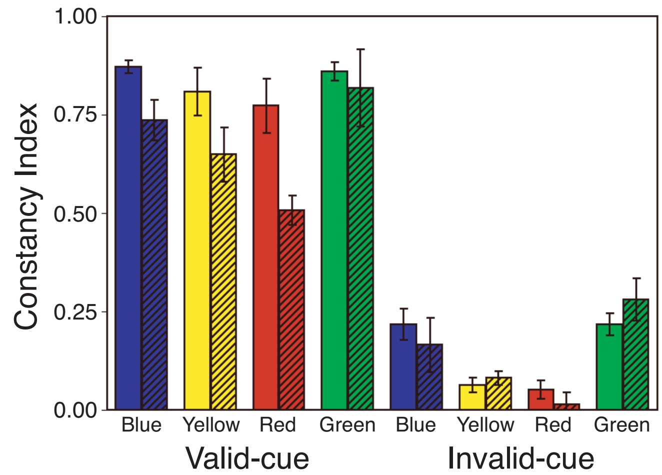

Figure 26. The constancy indices are shown for the decrements (solid bars) and increments (patterned bars) for the 30 ΔE* illuminant conditions. The error bars are +/– 1 SEM.

The recentering procedure in Step 1 is model-based. We wanted to know how sensitive the constancy index is to the choice of recentering procedure. Here we compare the results for two procedures: (a) the *diagonal procedure* (used for all constancy index calculations above) and (b) the *linear model procedure*.

In the diagonal procedure, the achromatic chromaticities measured for the reference illuminant and the experimental illuminant are used to derive a set of relative L-, M-, and S-cone gains that describe the difference in visual processing for tests embedded in the two experimental images. These derived gains are then applied to a stimulus with the chromaticity of the reference illuminant to produce the equivalent illuminants. This calculation is described in detail by Brainard (1998).

In the linear model procedure, three-dimensional models of surface reflectances (Nickerson, 1957) and illuminants (Judd et al., 1964) are used to calculate the surface reflectance that would produce the achromatic settings made by an observer under the standard illuminant. This surface reflectance is then used to calculate the equivalent illuminants required to produce the achromatic settings under the test illuminant conditions. The linear models are used to convert between chromaticities and surface and illuminant spectra. This model is closely related to the equivalent illuminant model used by Brainard et al. (1997), but here is applied to achromatic settings rather than to asymmetric matches.

For both procedures, CIs were calculated using both the Neutral and changed illuminant in the role of the reference illuminant and these were then averaged to provide the reported CI.

The CIs for Experiments 1 and 2 are shown in Figure 27 and the CIs for Experiment 3 are shown in Figure 28. The indices obtained by using the diagonal model procedure (solid bars) are compared with those obtained using

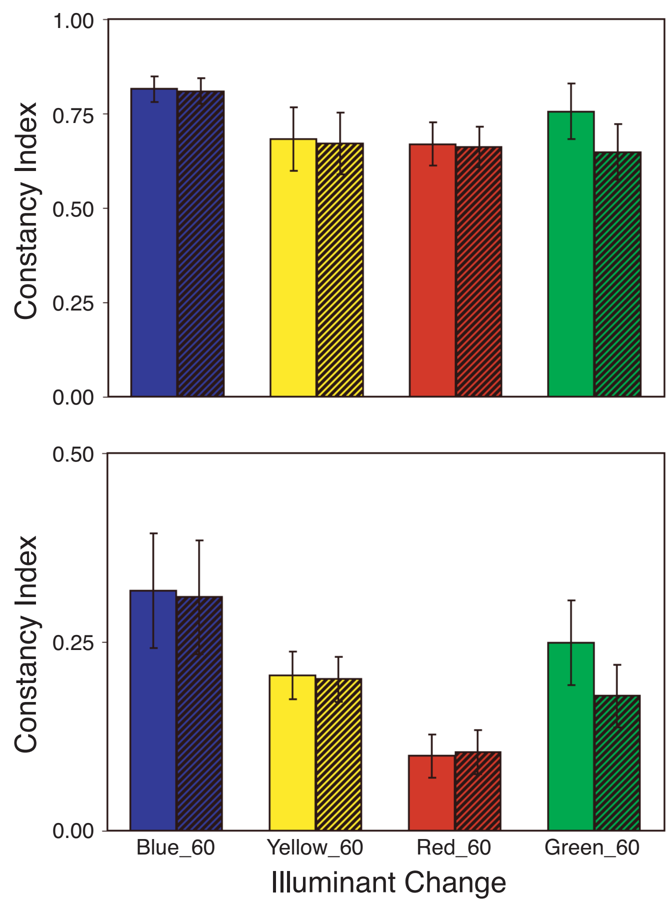

Figure 27. The constancy indices for the diagonal model (solid bars) and the linear model (striped bars) are shown for the 60 ΔE* illuminants for valid-cue conditions (top panel) and invalid-cue conditions (bottom panel). The error bars are +/– 1 SEM.

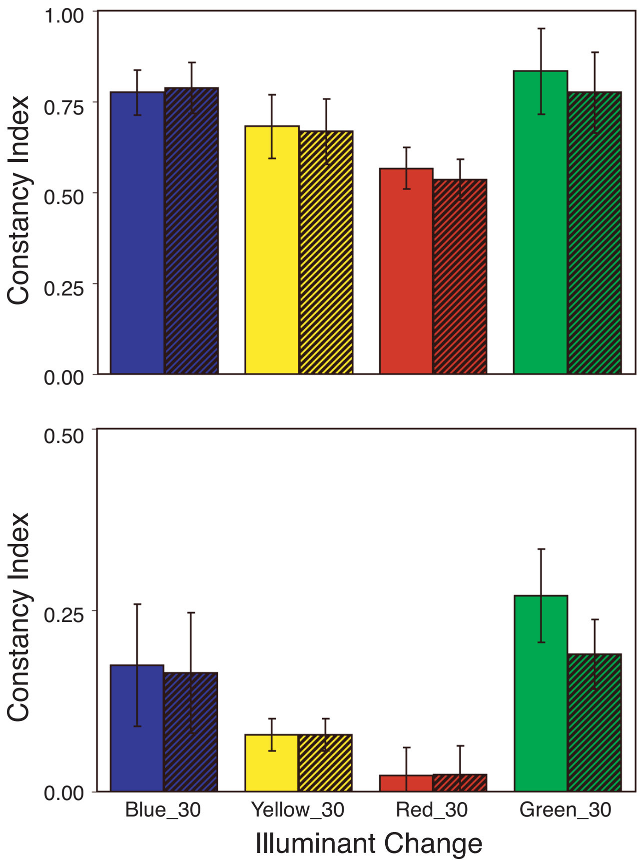

Figure 28. The constancy indices for the diagonal model (solid bars) and the linear model (striped bars) are shown for the 30 ΔE* illuminants for valid-cue conditions (top panel) and invalid-cue conditions (bottom panel). The error bars are +/– 1 SEM.

the linear model procedure (striped bars). The indices are similar in general, although the linear model procedure leads to a consistently lower index for Green illuminant changes. The differences between the indices do not lead to differences in the main conclusions we draw about how constancy depends on the color direction of the illuminant change.

Footnotes

[1]The simplified model assumes diffuse illumination and Lambertian surfaces. These assumptions are not met for real scenes, but the simplified model provides a useful starting point. See Foley, van Dam, Feiner, and Hughes (1990) for a description of a more elaborated model.

[2]Observers were given the following instruction: "While you're doing the experiment, it might be tempting to just assume that one of the areas you see is gray/white and then adjust the test patch until it looks like that area. It's very important that you do not do that. We want you to adjust the patch so that it looks gray/white, not so that it looks like some other area that you see."

[3]The recentering calculation is described in detail elsewhere (Brainard, 1998). Briefly, for each experimental illuminant, the achromatic chromaticities measured for the reference illuminant and the experimental illuminant are used to derive a set of relative L-, M-, and S-cone gains that describe the difference in visual processing for tests embedded in the two experimental images. These derived gains are then applied to a stimulus with the chromaticity of the reference illuminant to predict the chromaticity of the stimulus under the experimental illuminant that would have the same visual effect. Appendix C describes the index calculation in more detail.

[4]The 6-cm average separation is reported by Wandell (1995) and was provided to him by Ben Backus. Backus (personal communication) notes that 6 cm is rounded down from the average 6.25 cm interpupillary distance of his subjects.

References

Arend, L. E., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A, 3*, 1743-1751. [PubMed]

Arend, L. E., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A, 8*, 661-672. [PubMed]

Bauml, K. H. (1994). Color appearance: Effects of illuminant changes under different surface collections. *Journal of the Optical Society of America A, 11*, 531-542. [PubMed]

Bauml, K. H. (1999). Simultaneous color constancy: How surface color perception varies with the illuminant. *Vision Research, 39*, 1531-1550. [PubMed]

Bauml, K. H. (2001). Increments and decrements in color constancy. *Journal of the Optical Society of America A, 18*, 2419-2429. [PubMed]

Bloj, M., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402*, 877-879. [PubMed]

Brainard, D. H. (1989). Calibration of a computer controlled color monitor. *Color Research and Application, 14*, 23-34.

Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A, 15*, 307-325. [PubMed]

Brainard, D. H. (2003). Color constancy. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences*. Cambridge, MA: MIT Press.

Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). Color constancy in the nearly natural image. 1. Asymmetric matches. *Journal of the Optical Society of America A, 14*, 2091-2110. [PubMed]

Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A, 14*, 1393-1411. [PubMed]

Brainard, D. H., Kraft, J. M., & Longère, P. (2003). Color constancy: Developing empirical tests of computational models. In R. Mausfeld & D. Heyer (Eds.), *Colour perception: From light to object*. Oxford: Oxford University Press.

Brainard, D. H., & Wandell, B. A. (1991). A bilinear model of the illuminant's effect on color appearance. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing*. Cambridge, MA: MIT Press.

Brainard, D. H., & Wandell, B. A. (1992). Asymmetric color-matching: How color appearance depends on the illuminant. *Journal of the Optical Society of America A, 9*, 1433-1448. [PubMed]

Brainard, D. H., Wandell, B. A., & Cowan, W. B. (1989). Black light: How sensors filter spectral variation of the illuminant. *IEEE Transactions on Biomedical Engineering, 36*, 140-149. [PubMed]

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute, 310*, 1-26.

Burnham, R. W., Evans, R. M., & Newhall, S. M. (1957). Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America, 47*, 35-42.

Chichilnisky, E. J., & Wandell, B. A. (1996). Seeing gray through the on and off pathways. *Visual Neuroscience, 13*, 591-596. [PubMed]

Chichilnisky, E. J., & Wandell, B. A. (1999). Trichromatic opponent color classification. *Vision Research, 39*, 3444-3458. [PubMed]

CIE (1986). *Colorimetry* (2nd ed.). Paris: Bureau Central de la CIE.

D'Zmura, M., & Iverson, G. (1993). Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces. *Journal of the Optical Society of America A, 10*, 2148-2165. [PubMed]

D'Zmura, M., Iverson, G., & Singer, B. (1995). Probabilistic color constancy. In R. D. Luce, M. D'Zmura, D. Hoffman, G. Iverson, & A. K. Romney (Eds.), *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow's 70th Birthday* (pp. 187-202). Mahwah, NJ: Lawrence Erlbaum Associates.

D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A, 3*, 1662-1672. [PubMed]

Delahunt, P. B., & Brainard, D. H. (2000). Control of chromatic adaptation: Signals from separate cone classes interact. *Vision Research, 40*, 2885-2903. [PubMed]

DeMarco, P., Pokorny, J., & Smith, V. C. (1992). Full-spectrum cone sensitivity functions for X-chromosome-linked anomalous trichromats. *Journal of the Optical Society A, 9*, 1465-1476. [PubMed]

DiCarlo, J. M., & Wandell, B. A. (2000). *Illuminant estimation: Beyond the bases*. Paper presented at the Eighth Color Imaging Conference, Scottsdale, AZ.

Endler, J. A. (1993). The color of light in forests and its implications. *Ecological Monographs, 63*, 1-27.

Fairchild, M. D. (1990). *Chromatic adaptation and color appearance*. Unpublished doctoral dissertation, University of Rochester, Rochester.

Finlayson, G. D., Hubel, P. H., & Hordley, S. (1997). *Color by correlation*. Paper presented at the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems, and Applications, Scottsdale, AZ.

Foley, J. D., van Dam, A., Feiner, S. K., & Hughes, J. F. (1990). *Computer graphics: Principles and practice* (2nd ed.). Reading, MA: Addison-Wesley.

Forsyth, D. A. (1990). A novel algorithm for color constancy. *International Journal of Computer Vision, 5*, 5-36.

Foster, D. H., Amano, K., & Nascimento, S. M. (2003). Tritanopic color constancy under daylight changes? In J. D. Mollon, J. Pokorny, & K. Knoblauch (Eds.), *Normal and defective colour vision*. Oxford: Oxford University Press.

Funt, B. V., Drew, M. S., & Ho, J. (1991). Color constancy from mutual reflection. *International Journal of Computer Vision, 6*, 5-24.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science, 195*, 185. [PubMed]

Helson, H., & Michels, W. C. (1948). The effect of chromatic adaptation on achromaticity. *Journal of the Optical Society of America, 38*, 1025-1032.

Hochberg J. E., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology, 47*, 263. [PubMed]

Hurlbert, A. C. (1998). Computational models of color constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 283-322). Cambridge: Cambridge University Press.

Ishihara, S. (1977). *Tests for colour-blindness*. Tokyo: Kanehara Shuppen Company, Ltd.

Judd, D. B., MacAdam, D. L., & Wyszecki, G. W. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America, 54*, 1031-1040.

Kaiser, P. K., & Boynton, R. M. (1996). *Human color vision* (2nd ed.). Washington, D.C.: Optical Society of America.

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences, USA, 96*, 307-312. [PubMed]

Kraft, J. M., Maloney, S. I., & Brainard, D. H. (2002). Surface-illuminant ambiguity and color constancy: Effects of scene complexity and depth cues. *Perception, 31*, 247-263. [PubMed]

Langer, M. S., & Bülthof, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception, 29*, 649-660. [PubMed]

Larson, G. W., & Shakespeare, R. (1998). *Rendering with radiance: The art and science of lighting visualization.* San Francisco: Morgan Kaufman Publishers.

Lee, H. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A, 3*, 1694-1699.

Lucassen, M. P., & Walraven, J. (1996). Color constancy under natural and artificial illumination. *Vision Research, 36*, 2699-2711. [PubMed]

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In K. T. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 387-416). Cambridge: Cambridge University Press.

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

Mausfeld, R. (1998). Color perception: From Grassman codes to a dual code for object and illumination colors. In W. G. K. Backhaus, R. Kliegl, & J. S. Werner (Eds.), *Color vision - perspectives from different disciplines* (pp. 219-250). Berlin: Walter de Gruyter & Co.

Mausfeld, R., & Niederee, R. (1993). An inquiry into relational concepts of colour, based on incremental principles of colour coding for minimal relational stimuli. *Perception, 22*, 427-462. [PubMed]

Nickerson, D. (1957). *Spectrophotometric data for a collection of Munsell samples.* Washington, D.C.: U.S. Department of Agriculture.

Rutherford, M. D., & Brainard, D. H. (2002). Lightness constancy: A direct test of the illumination estimation hypothesis. *Psychological Science, 13*, 142-149. [PubMed]

Ruttiger, L., Mayser, H., Serey, L., & Sharpe, L. T. (2001). The color constancy of the red-green color blind. *Color Research and Applications, 26*, S209-S213.

Schirillo, J. A. (1999a). Surround articulation. I. Brightness judgments. *Journal of the Optical Society of America, 16*, 793-803. [PubMed]

Schirillo, J. A. (1999b). Surround articulation. II. Lightness judgments. *Journal of the Optical Society of America, 16*, 804-811. [PubMed]

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

Smith, V., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research, 15*, 161-171. [PubMed]

Speigle, J. M., & Brainard, D. H. (1999). Predicting color from gray: The relationship between achromatic adjustment and asymmetric matching. *Journal of the Optical Society of America A, 16*, 2370-2376. [PubMed]

Trussell, H. J., & Vrhel, M. J. (1991). *Estimation of illumination for color correction*. Paper presented at the International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada.

Wandell, B. A. (1995). *Foundations of vision*. Sunderland, MA: Sinauer.

Werner, J. S., & Walraven, J. (1982). Effect of chromatic adaptation on the achromatic locus: The role of contrast, luminance and background color. *Vision Research, 22*, 929-944. [PubMed]

Wyszecki, G., & Stiles, W. S. (1982). *Color science - concepts and methods, quantitative data and formulae* (2nd ed.). New York: John Wiley & Sons.

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: Tests of three candidate cues. *Vision Research, 41*, 2581-2600. [PubMed]