Color constancy in the nearly natural image. I. Asymmetric matches

David H. Brainard, Wendy A. Brunt, and Jon M. Speigle

Department of Psychology, University of California, Santa Barbara, Santa Barbara, California 93106

Received November 6, 1996; revised manuscript received March 4, 1997; accepted March 4, 1997

Most empirical work on color constancy is based on simple laboratory models of natural viewing conditions. These typically consist of spots seen against uniform backgrounds or computer simulations of flat surfaces seen under spatially uniform illumination. We report measurements made under more natural viewing conditions. The experiments were conducted in a room where the illumination was under computer control. Observers used a projection colorimeter to set asymmetric color matches across a spatial illumination gradient. Observers' matches can be described by either of two simple models. One model posits gain control in cone-specific pathways. This diagonal model may be linked to ideas about the action of early visual mechanisms. The other model posits that the observer estimates and corrects for changes in illumination but does so imperfectly. This equivalent illuminant model provides a link between human performance and computational models of color constancy. © 1997 Optical Society of America [S0740-3232(97)01609-8]

1. INTRODUCTION

An important function of vision is to provide information about objects. This task is difficult because there is no simple mapping between an object's intrinsic properties and the retinal image. Although the image depends on the object's properties, it is also affected by extrinsic factors, including where the object is located, how it is oriented, and how it is illuminated. To generate a representation that provides reliable information about objects, a visual system must compensate for these extrinsic factors.

In the case of object color, the spectral power distribution of the light reflected from an object depends on the object's intrinsic surface reflectance, on the illumination, and on the viewing geometry. The ability of a visual system to maintain object color appearance across variations in factors extrinsic to the object is called color constancy. It has long been clear both that human vision exhibits some degree of color constancy and that this color constancy is not prefect.[1,2]

Since humans are not perfectly color constant, a possible goal for experimentation is to determine how closely human performance approximates perfect constancy.[3–7] A more general goal is to develop principles and models that allow us to predict color appearance in complex scenes.[7–22] Ultimately, such models should account for color appearance in natural scenes.

Most empirical work on the effect of the viewing context is based on simple laboratory models of natural viewing conditions. These laboratory models typically consist of spots (or simple patterns) seen against uniform backgrounds[5,10,11,13,14,18,19,23–25] or computer simulations of flat matte surfaces seen under spatially uniform illumination.[3,6,7,15–17,26–30] Taken as a whole, these experiments reveal consistent regularities. First, changing the context changes the gains of the signals transmitted by the three classes of cones, an effect often referred to as von Kries adaptation.[8,15,18,19] Second, changing the context changes an additive contribution to the signals transmitted by the three classes of cones.[10,11,13,14,18,19] Although it is clear that color opponency plays an important role in the coding of cone signals by the nervous system,[24,31,32] this opponency does not manifest itself in most experiments designed to measure the effect of context on color appearance (but see Ref. 25). Similarly, although there must be nonlinearities as signals are transmitted from the cones to higher visual areas in the cortex, it is not generally necessary to model these nonlinearities to understand color context effects (but see Refs. 7, 30, and 33). One possibility is that opponency and the major nonlinearities occur at neural sites central to those where context exerts its influence.

There is no guarantee that regularities found for simple stimuli will generalize to more natural viewing conditions. A number of investigators have studied the color appearance of illuminated surfaces that are perceived as such[12,34–37] (see also Ref. 38). To the extent that they speak to the same issues, these studies are generally consistent with the regularities revealed by experiments that use simpler stimuli. In the lightness domain, however, Gilchrist and his colleagues[39–42] have shown that how the perceptual system parses a complex scene into surfaces and illuminants can have a large effect on perceived lightness (see also Refs. 43 and 44). Their results suggest a richness to color context that has not yet been extensively explored.

This paper presents experiments designed to extend our understanding of color constancy. To perform our experiments, we have constructed and calibrated an experimental room where the illumination is under computer control. An apparatus in the room allows us to measure asymmetric color matches[10,15,18,26,45] under nearly natural viewing conditions. Our emphasis is on developing models that characterize the effect of the illuminant on color appearance.

In natural scenes there are two quite distinct ways that the illumination can vary. First, the illumination can change from one time to another. For example, the spectrum of the light at dawn tends to differ from that at noon. Second, the illumination can change from one location to another within a single scene. For example, the spectrum of the illumination in shadows differs from that of direct sunlight. We use the term successive color constancy to refer to constancy with respect to illumination changes that occur over time and the term simultaneous color constancy to refer to constancy with respect to changes that occur across space. The literature does not generally distinguish between these two types of constancy, and it is possible that the two are mediated by distinct visual mechanisms. Indeed, computational analyses of successive color constancy[46–48] suggest that better estimates of the illuminant may be obtained if the visual system integrates information from many surfaces, whereas simultaneous constancy requires that this integration not extend across illumination boundaries. Moreover, different considerations may guide experimental study of the two types of constancy. In studying successive constancy, it is desirable to control the observer's state of adaptation carefully.[10,15] In studying simultaneous color constancy, on the other hand, an ecologically valid task is for the observer to compare the colors of two objects, each seen under its own illumination. In this case, manipulations designed to achieve precise control of adaptation may make the observer's task difficult or unnatural. Here we emphasize that we are presenting experiments designed to measure simultaneous color constancy.

2. GENERAL METHODS

A. Overview

To study color constancy across illumination gradients, our general strategy was to have observers make asymmetric color matches. Observers viewed colored surfaces in an experimental room where the illumination was under computer control (Apple Macintosh Quadra 700). The observers' task was to adjust the appearance of a match surface, seen on one side of the room, to be the same as that of a test surface, seen on the other side. Observers adjusted appearance by controlling the output of a colorimeter that projected directly onto the match surface. We had observers make asymmetric matches to multiple test surfaces under different illumination conditions. In control conditions the illumination was spatially uniform, so that the test and match surfaces were illuminated identically. In experimental conditions we introduced a spatial gradient of illumination so that the test and match surfaces were illuminated differently.

B. Experimental Room

Our experimental room is shown schematically in Fig. 1. The ambient illumination of the room was produced by three sets of computer-controlled stage lamps (SLD Lighting, 6-in. (15-cm) Fresnel #3053, BTL 500-W bulb) arranged in four triads as shown in the figure. The set labeled R had red dichroic filters (Rosco 6100 "Flame Red"); the set labeled G had green dichroic filters (Rosco 4959 "Light Green"); the set labeled B had blue dichroic filters

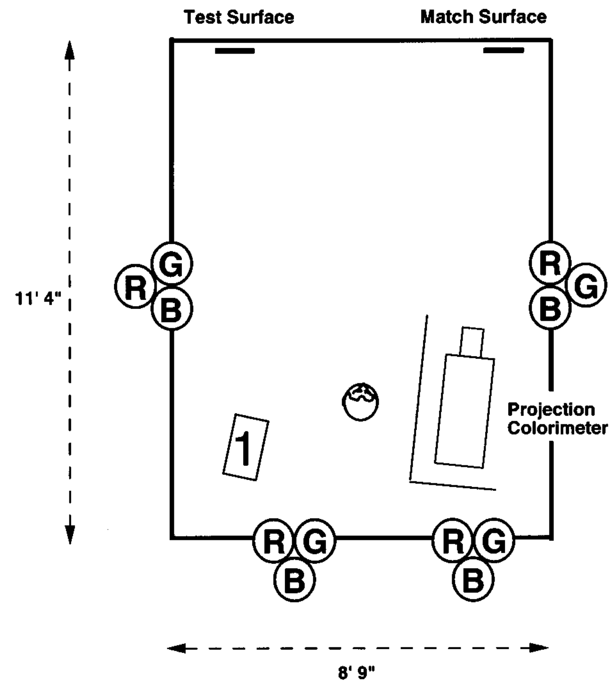

Fig. 1. Experimental room. The room is 8′9″ × 11′4″. Its walls and ceiling were painted a matte gray of roughly 50% reflectance; its floor was covered with a gray carpet. Other objects in the room were visible to the observer, including a brown metal bookcase and a white table. Ambient illumination was provided by 12 theater stage lamps, labeled R, G, and B, arranged in four triads. Gradients of illumination were provided by an additional lamp, indicated by the rectangle labeled 1. The observer viewed test and match surfaces located on the far wall of the room. The match surface was spot illuminated by a projection colorimeter. Both test and match surfaces were surrounded by a 1/4″ black felt border. The background for the surfaces was a 48″ × 72″ sheet of particle board painted the same gray as the room. Locations indicated by the figure are approximate.

(Rosco 4600 "Blue"). The light from each triad was passed through a gelatin diffuser to minimize colored shadows. The lamps were controlled from software by varying the rms voltage across the bulbs (NSI 5600 Dimmer Packs, NSI OPT-232 interface card, 100 voltage quantization levels). We yoked the voltages of all lights within a set (R, G, or B) together. By varying the intensities of the three sets of lamps, we varied the spectral power distribution of the ambient illumination. Control software corrected for spectral shifts introduced when the voltage to the bulbs was varied. A spatial gradient of illumination was provided by an additional lamp placed at the location labeled 1 in the figure.

Test and match surfaces were 8.5-in. (21.25-cm) by 11 in. (27.5-cm) matte Munsell papers placed on the far wall of the room. Each surface was surrounded by a 0.25-in. (0.625-cm) black felt border. From the observer's point of view the test surface occupied 4.9° horizontal by 6.4° of visual angle, while the match surface occupied 4.7° by 6.1°. The separation between the test and match surfaces was 50-in. (125-cm) or 25° of visual angle. The immediate background for both surfaces was a 48-in. (120-cm) by 72-in. (180-cm) sheet of particle board painted a matte gray of roughly 50% reflectance.

Light was reflected from the test surface to the observer in the normal manner. The light reflected from the match surface, however, consisted of two components. The first was normal reflection of the ambient illumination. The second was generated by a computer-controlled projection colorimeter (see figure). The illumination from the colorimeter was masked so that it was spatially coincident with the match surface.

The colorimeter had three independently controlled primaries. The light source for each primary was a slide projector bulb (Type FHS, 300 W, 82V). Light from each bulb passed through a heat-absorbing filter and a red, green, or blue dichroic filter (OCLI). The light from the three bulbs was then combined with dichroic beam splitters (OCLI) and passed through a slide projector condenser (Kodak 4400), an adjustable mask, and a slide projector lens (WIKO, 100 mm, *f* 2.8). The intensity of each primary was controlled by adjusting the rms voltage supplied to the corresponding projector lamp (NSI 5600 Dimmer Packs, NSI OPT-232 interface card, factory modified to provide 256 voltage quantization levels, power resistor used to drop voltage range to 0–82 V). Control software corrected for spectral shifts in the primaries introduced by the variation in the rms voltage to the bulbs.

The use of the colorimeter allowed us to manipulate the tristimulus coordinates of the light reaching the observer from the match surface. For the conditions reported here, the match surface appeared to be a normal surface. The illumination from the colorimeter was not explicitly visible; varying it changed the apparent color of the match surface. This subjective observation is confirmed indirectly by control conditions described below.

C. Experimental Procedure

During an experimental session, the observer made asymmetric matches to a number of different test surfaces under a single illumination condition. The observer (Experiments 1, 2a, and 2b observer JMS) or an experimenter (Experiment 2b observers ASH and PBE) began a match by mounting the desired test surface on the far wall of the room. The gamut of possible match tristimulus coordinates depended on both the illumination and the identity of the match surface. To achieve the broadest possible gamut for our experiments, we varied the match surface from condition to condition. We used pilot data to determine which match surfaces would allow the observer to make a match. For each test surface and illumination condition, then, the observer mounted the match surface that was most likely, given pilot data, to allow a match to be made.

Once the surfaces were mounted, the observer adjusted the colorimeter to make the match surface appear as similar as possible to the test surface. The observer was free to look back and forth at the two surfaces while setting the match. Indeed, the two surfaces were far enough apart so that it was difficult to compare them simultaneously. The observer controlled the appearance of the match surface by using three knobs (optical shaft encoders interfaced through custom hardware to a National Instruments LAB-NB interface card). In Experiments 1 and 2a, each knob varied an approximation to the three coordinates of the CIELAB uniform color space.[49] These coordinates correspond to the lightness, red–green, and blue–yellow perceptual dimensions of color experience. In addition, the space is approximately uniform, so that equal steps along each coordinate axis tend to have equal perceptual effects. In Experiment 2b, the three knobs directly varied the output of one of the three colorimeter primaries. If the observer was not able to make a satisfactory match, he or she was free to try another match surface to move the colorimeter gamut.

When the observer had set the best possible match, he or she indicated its quality (good, OK, or bad for Experiments 1 and 2a; good or bad for Experiment 2b). The observer then made direct spectral measurements (Photo-Research PR-650) of the light reflected from both the test and the match surfaces. The data that we report are the results of these measurements and specify the proximal stimulus reaching the observer. This measurement procedure ensures that our data are not corrupted by any inaccuracies in the characterization and control of the illuminants and the colorimeter. The results of the measurements were recorded directly by the control computer, so that the observer received no feedback during the session about the physical properties of his or her matches.

D. Instructions

Observers were instructed to adjust the match surface so that it had the same color appearance as the test surface. In simultaneous matching experiments performed on a CRT monitor, Arend and colleagues[26,27] found that observers were able to perform two distinct color matching tasks. In the unasserted-color task (also called the hue-matching task), observers were asked to match the color sensation produced by the test and match stimuli. In the perceived-surface-color task (also called the paper-matching task), observers were asked to make the test and match stimuli appear as if they had the same surface reflectance. Kuriki and Uchikawa[50] and Bauml[51] recently confirmed that observers are able to perform these two tasks under certain experimental conditions. In these experiments, the test and match surfaces are simulated in the context of identical surround surfaces. Bauml[52] describes the perceived-surface-color task as involving relational judgments; the observer proceeds (at least in part) by adjusting the match surface so that the color relations between it and its surround are similar to those between the test surface and its surround.

In pilot experiments, authors DHB and WAB tried to distinguish the unasserted-color and perceived-surface-color tasks for our viewing conditions. In contrast to the experiments of Arend and colleagues[26,27] and of Bauml,[51] our test and match stimuli are not placed in identical multicolored surrounds, although there is a common neutral background near both the test and the match surfaces. We found the unasserted-color task straightforward but had difficulty distinguishing a separate perceived-surface-color task. Our stimuli look like surfaces, and our illumination gradient is gradual; we found that the color sensation produced by our stimuli and their identity as surfaces were closely intertwined. We did not explicitly try a relational strategy, as our conversations with Bauml took place after we conducted our pilot experiments. Because we had difficulty distinguishing the two tasks, we were not confident that we could intelligently instruct naïve observers to perform one or the other. We therefore simply asked our observers to make color matches. We believe that in our experiments these instructions lead to judgments similar to those of Arend and colleagues[26,27] for the unasserted-color task.

E. Control of Illumination

Our illuminant-control procedure began with a characterization of the four RGB triads of lamps that provided the ambient illumination of the room. This characterization, described below, allowed us to compute device settings to achieve any desired tristimulus coordinates at the test surface location in the absence of other light sources. In some experimental conditions, however, the gradient-producing lamp at location 1 of Fig. 1 was on, and it affected the illumination at the test surface as well as at the match surface. In these conditions, we measured and compensated for this contribution to the test illuminant when we set the ambient illumination. At the start of each session, we made direct measurements of the actual illuminant at the test and match surface locations (PhotoResearch PR-650, RS-2 reflectance standard). The data that we report are based on these measurements.

Our procedure for characterizing the RGB illuminant triads was as follows. For each triad, we measured the full spectral power distribution of the illuminant at the test surface for 25 evenly spaced device settings. The measurements were made with a spectraradiometer (PhotoResearch PR-650), which imaged a reflectance standard (PhotoResearch RS-2) placed directly in front of the test surface. These and all 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[49] of 5-nm increments between 380 and 780 nm. We fitted the measured spectra with a two-dimensional linear model.[53] Two-dimensional linear models provided good fits to the measured spectral data, accounting for essentially all of the measured variance. For each set, the individual spectra could thus be represented by using two coefficients. The first coefficient specified the contribution of the principal component of the entire data set to the spectrum. The second component specified a smaller correction that accounted for shifts in the lamp spectra as the rms driving voltage was varied. We interpolated the measured relations between control voltage and linear model coefficients to provide tables that specified the linear-model coefficients corresponding to each possible device setting. Together with the spectra of the linear-model basis functions, this table allowed us to predict the illuminant spectra produced by any device setting.

Our control software used the characterization table and the basis functions described above. To produce an illuminant with given tristimulus coordinates, we began by neglecting the higher-order term in the linear model. In this case we could calculate the desired linear-model coefficients by using standard methods.[53] We then searched our tables to find the device settings for the R, G, and B triads that came closest to producing these coefficients. This initial procedure ignores the higher-order components of the linear model. Given the initial device settings, however, we could use the full table to predict what illuminant tristimulus values would actually be produced. We used the difference between predicted and desired tristimulus values to drive an iterative search procedure. Although we were not guaranteed that our search procedure would always converge on the best possible values, in practice we found that it gave good performance. As noted above, we report data based on actual measurements of the illuminants so that any error in our control procedure is not propagated to our data.

F. Colorimeter Control

Because our experiments employ a matching procedure in which the observer adjusts the colorimeter to achieve a desired appearance, precise characterization and control of the projection colorimeter was not required. As noted above, we made direct measurements of the tristimulus coordinates of the test and match surfaces after each match. For Experiments 1 and 2a we tried to arrange the knobs so that they had an intuitive perceptual effect. We characterized the projection colorimeter by using the same general procedures as described above for characterizing the illuminant; we did this characterization in the dark while using a Munsell N 3/ paper as the match surface. Knob settings were interpreted as L^*, a^*, and b^* coordinates of the CIELAB uniform color space and were converted by software to tristimulus coordinates. These were then converted to device settings with a single iteration of the procedure described above for controlling the illuminant. Since we did not adjust the calibration data for different match surfaces or different ambient illuminants, the correspondence between the three knobs and the L^*, a^*, and b^* coordinates was only approximate. For Experiment 2b we simply let the each knob directly control the output of one of the colorimeter primaries. Observers were able to set matches with either control scheme.

3. EXPERIMENT 1: CONTROL

A. Rationale

Experiment 1 had two purposes. The first was to validate our general experimental procedure. For our results to be valid, it is important that the tristimulus coordinates of observers' matches be independent of the identity of the match surface mounted at the match location. A dependence would indicate that the observer was separating the two components of the light reflected from the match surface (one from the ambient and one from the colorimeter). The second purpose was to establish how well observers could set matches when no illumination gradients were introduced in the room. We wanted to be sure our task was possible in a control condition before trying to use it to study the effect of illumination gradients.

B. Methods: Experiment 1

The methods used in this experiment are as described in Section 2. We used no gradient-producing lamps, so the ambient illumination was approximately uniform throughout the room and very similar at the test and match surface locations. Thus the observer's task was quite similar to that used in the classic color matching experiment. The main difference is that in our experiment the test and match surfaces were spatially separated. Their immediate surrounds, however, were nearly identical. The measured chromaticities and luminance of the test illuminant were $x = 0.309$, $y = 0.348$, $Y = 21.15$ cd/m^2. The measured chromaticities and luminance of the match illuminant were $x = 0.299$, $y = 0.329$, $Y = 19.24$ cd/m^2. The difference between test and match illuminants is indicative of the spatial uniformity of the illumination across the front of the room.

We had observers set color matches to a number of different test surfaces. We report data only for those test surfaces for which observers indicated that the match was good. Matches were set in several experimental sessions. A typical session lasted about 1 h, and an observer could typically set between eight and ten matches in this amount of time. Observers made between one and seven matches for each test-surface/match-surface pair.

Two observers participated in Experiment 1. They were the second author (WAB, age 22) and a paid undergraduate (TES, age 20). Both observers were female and presumably color normal. Observer TES was naïve as to the purpose of the experiment.

C. Results: Experiment 1

The results of Experiment 1 validate our general experimental design and procedures. Figure 2 shows the results for two test surfaces for observer WAB. The open squares show the coordinates of the test surfaces, and the solid squares show the coordinates of matches made with different match surfaces. Figure 2(a) is a chromaticity plot, and Fig. 2(b) plots luminance against x chromaticity. Two facts are evident from the figure. First, the matches are independent of which surface is mounted at the matching location. As noted above, this is important because it verifies that the observer is not sensitive to the

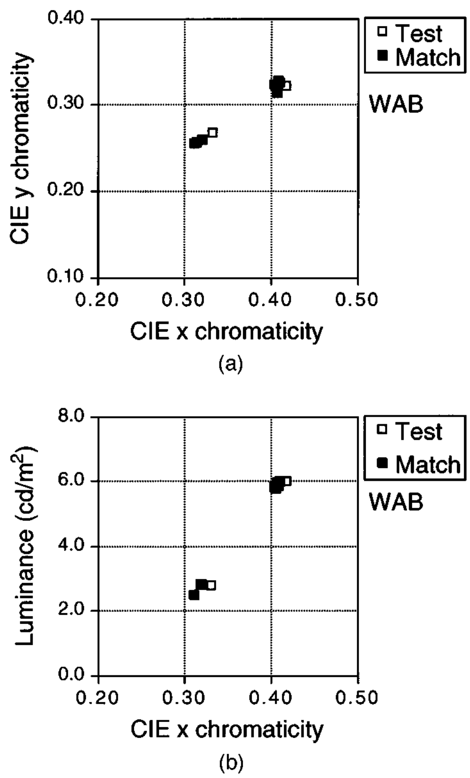

Fig. 2. Typical results from Experiment 1. (a) CIE xy chromaticity plot, (b) luminance versus x chromaticity. Open squares, coordinates of two test surfaces; solid squares, coordinates of matches made with different match surfaces. Each plotted point is the mean over replications for the same match surface.

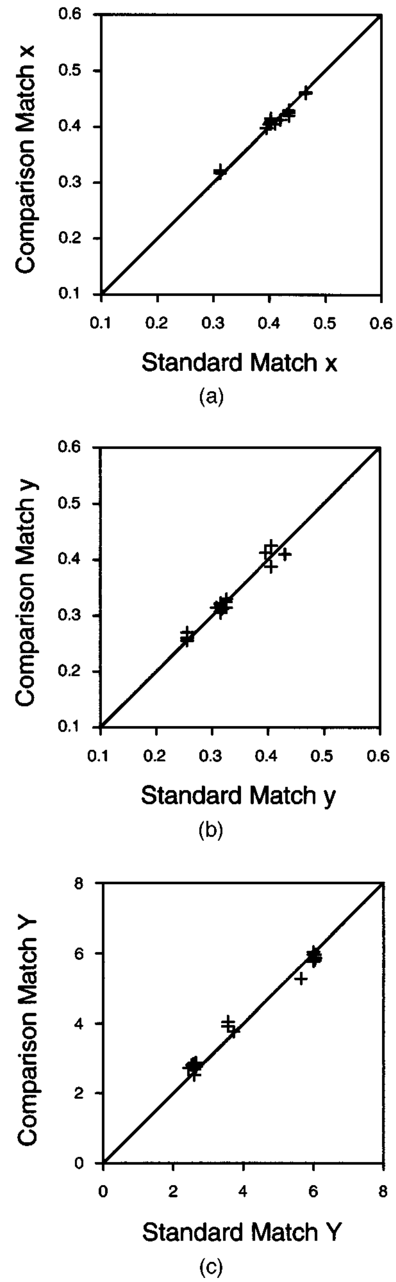

Fig. 3. Effect of match surface, Experiment 1. The data show that changing the identity of the match surface does not affect observers' matches. See description in text. (a) CIE x chromaticities of the comparison matches plotted against the CIE x chromaticity of the standard match; (b) CIE y chromaticities of the comparison matches plotted against the CIE y chromaticity of the standard match; (c) luminances of the comparison matches plotted against the luminance of the standard match. The data are collapsed across observers WAB and TES.

identity of the match surface. The close clustering of the solid squares also indicates that observers can perform the matching task reliably. Second, the matches are close to colorimetric matches. Although this is not surprising, it provides important verification that observers are able to perform the matching task under our conditions.

Figure 3 provides a more complete check that changing the identity of the match surface does not affect observers' matches. For each test surface for which multiple match surfaces were used, we chose the mean match from one match surface as a standard. The standard was based on the match surface for which we had the largest number of replications. We designated the mean matches from the other match surfaces as comparison matches. Figure 3(a) plots, for each test surface, the CIE x chromaticities of the comparison matches against the CIE x chromaticity of the standard match. If the identity of the matching surface has no effect on the observers' matches, then the data should fall along a line with slope 1. In fact, the data lie very close to this line. Similar results are seen in Figs. 3(b) and 3(c), which show corresponding plots for CIE y chromaticity and luminance.

4. EXPERIMENT 2: EFFECT OF AN ILLUMINATION GRADIENT

A. Rationale

The purpose of Experiment 2 was to investigate the effect of introducing a spatial illumination gradient. In natural scenes, a number of different processes may cause such gradients. For example, if the illumination is formed by a combination of diffuse blue skylight and a more directional yellow light from the disk of the sun, then both the chromaticity and the luminance of the overall illuminant may vary gradually across the scene as the relative contributions of these two components varies. On the other hand, shadows caused by occlusion may produce more abrupt changes in the relative contributions of the two components. In this experiment, the illuminant varied gradually between the test surface and the match surfaces.

B. Method: Experiment 2a

The methods used in this experiment are essentially the same as in Experiment 1, except that a gradual spatial illumination gradient was added to the ambient illumination by use of the gradient lamp (SLD Lighting, 6-in. (15-cm) Fresnel #3053, BTL 500-W bulb). The lamp had a yellow gelatin filter (Roscolux #08) This spatial gradient of illumination was arranged so that the test and match surfaces were illuminated quite differently. The illuminant was considerably more bluish at the test location than at the match location.

As in Experiment 1, we had observers set color matches to a number of different test surfaces. Again, we report data for only those test surfaces for which observers indicated that the match was good. Observers made between one and three matches for each test-surface/match-surface pair.

The observers were instructed to adjust the color of the match surface so that it appeared to look the same as the test surface. They were told that the illumination was different at the test and the match surface locations. They were not given any instructions about how the illumination gradient might affect their match settings. The chromaticities and luminances of the test and match illuminants are given in Table 3 below. The relevant columns are those for observers WAB and TES.

The same two observers as in Experiment 1 participated in this experiment.

C. Results: Experiment 2a

Introducing a spatial illumination gradient had a marked effect on the asymmetric matches. Figure 4 shows the results for both observers. The top panels show the results for observer WAB, and the bottom panels show the results for observer TES. The open squares show the coordinates of the test surfaces, and the solid squares show the coordinates of the match surfaces. Each corresponding pair of test and match surfaces is connected by a solid line. The standard errors for the matches are typically smaller than the plotted points. Table 1 provides the data from Experiment 2a in numerical form.

As in Experiment 1, we wanted to determine whether the matches were independent of which surface is mounted at the matching location. Figure 5 provides another check that changing the identity of the match surface does not affect observers' matches. This figure is in the same format as Fig. 3, which summarized similar data for Experiment 1. As with Experiment 1, the figure shows that observers' matches are independent of the identity of the match surface. Because the identity of the match surface does not seem to affect the matches, Table 1 presents the data collapsed across match surfaces.

D. Methods: Experiment 2b

Experiment 2b was similar to Experiment 2a. The main difference was that in Experiment 2b the test and match surfaces were seen in the context of 13 other 8.5-in. (21.25-cm) by 11-in. (27.5-cm) Munsell papers rather than against the uniform gray particle board background of Experiment 2a. These papers, along with the test and match surfaces, were arranged in a rectangular grid against the particle board background. There were three grid rows, with each row containing five surfaces. The leftmost surface in the center row was the test surface. The rightmost surface in the center row was the match surface. There are conflicting reports in the literature about how the number of surfaces in a scene affects color constancy.[26,29,37,50] These reports are difficult to reconcile at present, because different labs have used different methods and studied different aspects of constancy. Because the viewing conditions of Experiment 2a were already quite rich, we do not expect that our manipulation will settle this issue. Nonetheless, we thought it worthwhile to see whether adding more surfaces affected performance.

The illuminants used in Experiment 2b also differed slightly from one another and from those used in Experiment 2a. The chromaticities and luminances of the test and match illuminants for each observer are given in Table 3 below. The relevant columns are those for observers JMS, ASH, and PBE.

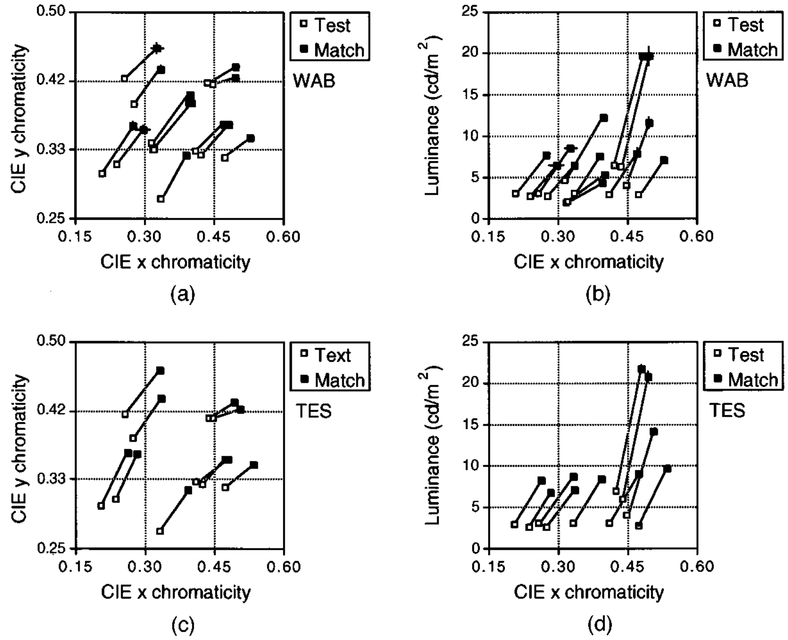

Fig. 4. Results from Experiment 2a. (a), (b) Results for observer WAB; (c), (d) results for observer TES. Open squares, test surface coordinates; solid squares, match surface coordinates. Each corresponding pair of test and match surfaces is connected by a solid line. For each test surface, each plotted point was obtained by taking the mean match over replications and over match surfaces. Where visible, the error bars represent ±1 standard error of the mean. Error bars not visible are smaller than the plotted points.

Table 1. Data for Experiment 2a, Observers WAB and TES[a]

Observer	Test Panel	Test x	Test y	Test Y	Match x	Match y	Match Y	Const. x	Const. y	Const. Y
WAB	5 R 6/6	0.421	0.328	6.50	0.482	0.364	19.57	0.520	0.373	29.69
	5 R 4/6	0.472	0.325	2.85	0.529	0.349	7.05	0.557	0.355	12.07
	5 R 4/4	0.409	0.332	2.94	0.470	0.365	7.96	0.522	0.367	11.19
	5 Y 6/6	0.436	0.414	6.28	0.496	0.434	19.64	0.511	0.429	26.15
	5 Y 5/6	0.446	0.412	4.08	0.494	0.421	11.54	0.518	0.432	17.08
	5 G 4/6	0.257	0.420	3.10	0.327	0.457	8.56	0.309	0.479	7.76
	5 G 4/4	0.277	0.389	2.76	0.334	0.431	6.53	0.367	0.440	8.92
	5 B 4/6	0.206	0.305	3.05	0.273	0.363	7.76	0.278	0.365	8.72
	5 B 4/4	0.239	0.317	2.70	0.296	0.359	6.44	0.334	0.384	9.12
	5 P 4/6	0.335	0.275	3.00	0.389	0.327	7.56	0.444	0.326	10.02
	N 3/	0.316	0.334	1.86	0.394	0.391	4.36	0.424	0.385	4.34
	N 3.5/	0.318	0.334	2.08	0.400	0.391	5.28	0.433	0.391	7.15
	N 5/	0.315	0.343	4.61	0.397	0.401	12.26	0.431	0.392	15.74
TES	5 R 6/6	0.422	0.328	6.86	0.477	0.359	21.79	0.520	0.373	29.69
	5 R 4/6	0.473	0.324	2.77	0.533	0.351	9.66	0.557	0.355	12.07
	5 R 4/4	0.410	0.331	2.99	0.473	0.358	9.03	0.522	0.367	11.19
	5 Y 6/6	0.437	0.407	5.89	0.493	0.427	20.78	0.511	0.429	26.15
	5 Y 5/6	0.448	0.409	3.99	0.504	0.419	14.13	0.518	0.432	17.08
	5 G 4/6	0.256	0.412	3.08	0.330	0.465	8.76	0.309	0.479	7.76
	5 G 4/4	0.273	0.383	2.56	0.333	0.433	7.10	0.367	0.440	8.92
	5 B 4/6	0.204	0.302	2.94	0.260	0.365	8.22	0.278	0.365	8.72
	5 B 4/4	0.236	0.310	2.57	0.281	0.365	6.84	0.334	0.384	9.12
	5 P 4/6	0.331	0.272	2.99	0.390	0.320	8.46	0.444	0.326	10.02

[a] The table provides the Munsell designations for the test surfaces, chromaticities and luminances of the test surfaces under the test illuminant, and chromaticities and luminances of the observers' asymmetric matches. Since the data indicate that there is no effect of match surface on the observers' matches (see Figs. 2, 3, and 5), data are collapsed across match surfaces. The final three columns of the table provide a color constancy prediction based on measurements of the light reflected from the test surface when it was placed at the match location (with the colorimeter turned off). Chromaticities are 1931 CIE x and y. Luminances are cd/m^2.

Three observers, JMS (male, age 29), ASH (female, age 27), and PBE (male, age 33) participated in Experiment 2b. Observer JMS is one of the authors; observers ASH and PBE were naïve. Observers JMS and PBE were color normal as tested by the Ishihara plates. In addition, JMS sets normal Rayleigh matches.

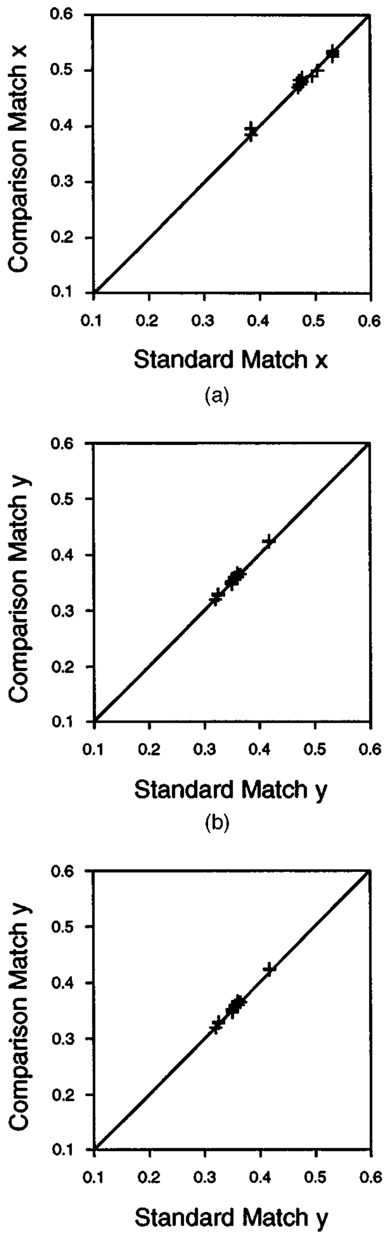

Fig. 5. Effect of match surface, Experiment 2a. The data again show that changing the identity of the match surface does not affect observers' matches. See description in text. The format and observers are the same as for Fig. 3.

E. Results: Experiment 2b

The results from Experiment 2b were very similar to those from Experiment 2a. Figure 6 shows the data for observer JMS, while Table 2 tabulates the data for all three observers. In Subsection 6.A we compare quantitatively the degree of constancy shown in Experiments 2a and 2b.

F. Discussion: Experiment 2

1. Quality of the Matches

Although the observers' tasks were identical in Experiments 1 and 2a, both observers reported that the quality of matches they could make in Experiment 2 was lower than that generally obtained in Experiment 1. Most matches were still rated as "good" in Experiment 2, but observers noted that this equivalence was the result of a criterion shift. They assigned a rating of "good" to the best matches within the second experiment without trying to equate the quality of these matches to those obtained in Experiment 1.

The observers were able to set reliably what they regarded as the best match. At this match point, however, the test and the match surfaces looked different, and the observers felt as if further adjustments of the match surface should produce a better correspondence. Yet turning any of the knobs or combinations of knobs only increased the perceptual difference. We verified that the observers' adjustments near the best match were not limited by the gamut of our apparatus.

It is somewhat difficult to express in words what is deficient about the matches set in Experiment 2. One might, however, say something like the following. At the best match, the test surface (seen under a bluish illuminant) has something of a cool cast about it, whereas the match surface (seen under a yellowish illuminant) has a warm cast. To the observer it seems therefore as if the match surface should be adjusted to be more bluish. But this adjustment does not change the warmth of the match surface. Rather, it has the effect of changing (say) a warm gray to a warm blue, which then still fails to match the cool gray test surface. Previous authors have attempted to describe and characterize how the quality of color appearance can vary across a variety of viewing conditions.[54–57]

In pilot experiments we set matches across a sharp illumination boundary introduced with spotlight (SLD Lighting, 4.5-in. (11.25-cm) Zoom Ellipsoidal #3062, EVR 500-W bulb, Roscolux #08 yellow filter). The chromaticities and luminances of the test and match illuminants were similar to those of Experiment 2a. Observers found setting satisfactory matches across this sharp illumination boundary very difficult and were successful for only a few test surfaces.

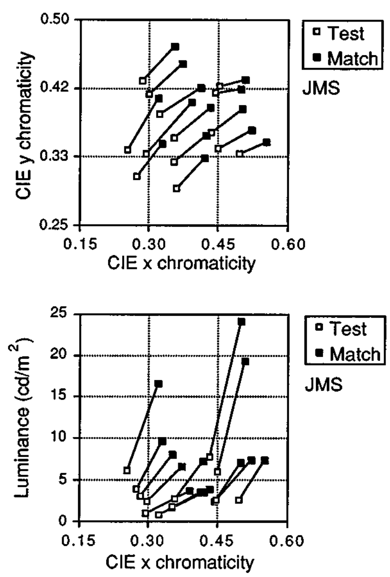

Fig. 6. Results from Experiment 2b. To avoid overwhelming the plots, only a subset of the data for observer JMS is shown. The format is as for Fig. 4.

Table 2. Data for Experiment 2b, Observers JMS, ASH, and PBE[a]

Observer	Test Panel	Test x	Test y	Test Y	Match x	Match y	Match Y	Const. x	Const. y	Const. Y
JMS	N 2.75/	0.351	0.353	1.20	0.434	0.392	3.02	0.458	0.389	5.22
	N 3/	0.354	0.358	1.71	0.432	0.393	3.93	0.461	0.391	7.42
	N 3.5/	0.356	0.354	1.83	0.435	0.391	4.57	0.460	0.392	7.93
	N 5/	0.352	0.353	4.08	0.424	0.395	10.11	0.456	0.392	17.62
	5 B 2.5/2	0.295	0.337	1.11	0.391	0.400	3.80	0.406	0.381	4.37
	5 B 3/2	0.302	0.346	1.60	0.378	0.408	4.23	0.411	0.390	6.39
	5 B 4/4	0.258	0.343	2.49	0.330	0.402	5.92	0.364	0.392	9.34
	5 B 6/6	0.254	0.341	6.20	0.320	0.405	16.62	0.360	0.391	23.18
	5 G 2.5/2	0.323	0.386	0.98	0.413	0.418	3.62	0.428	0.416	3.99
	5 G 3/2	0.327	0.379	1.46	0.415	0.424	3.79	0.434	0.410	6.08
	5 G 3/4	0.297	0.411	1.40	0.385	0.445	4.75	0.403	0.434	5.37
	5 G 4/4	0.299	0.409	2.44	0.373	0.446	6.67	0.406	0.433	9.66
	5 G 4/6	0.285	0.426	3.04	0.354	0.468	8.14	0.389	0.448	11.53
	5 G 6/6	0.285	0.422	5.84	0.357	0.473	16.93	0.389	0.448	22.94
	5 Y 3/1	0.380	0.374	1.40	0.451	0.398	3.36	0.480	0.401	6.23
	5 Y 4/4	0.444	0.411	2.43	0.499	0.415	7.21	0.520	0.417	11.52
	5 Y 5/6	0.462	0.420	3.99	0.508	0.420	10.77	0.529	0.420	19.15
	5 Y 6/6	0.451	0.420	6.03	0.508	0.427	19.39	0.522	0.422	28.61
	2.2 YR 6.47/4.12	0.433	0.363	7.75	0.501	0.392	24.13	0.525	0.381	35.61
	3 YR 3.7/3.2	0.457	0.366	2.24	0.515	0.380	6.60	0.538	0.383	10.63
	5 R 2.5/2	0.421	0.346	1.07	0.500	0.370	3.44	0.514	0.376	4.92
	5 R 3/2	0.414	0.350	1.71	0.492	0.375	4.46	0.508	0.380	7.87
	5 R 3/4	0.476	0.339	1.57	0.530	0.361	5.25	0.555	0.363	7.54
	5 R 4/4	0.449	0.344	2.72	0.522	0.366	7.52	0.537	0.371	13.22
	5 R 4/6	0.496	0.337	2.70	0.553	0.352	7.41	0.568	0.360	13.31
	5 R 6/6	0.455	0.345	6.28	0.522	0.369	21.07	0.536	0.376	30.82
	5 P 3/2	0.353	0.327	1.76	0.423	0.359	3.59	0.464	0.369	7.63
	5 P 3/4	0.359	0.303	1.62	0.429	0.347	4.76	0.468	0.352	6.99
	5 P 4/6	0.359	0.295	2.87	0.420	0.333	7.37	0.473	0.343	12.47
	5 P 6/6	0.352	0.310	6.55	0.422	0.354	18.95	0.462	0.358	28.10
	4.3 PB 4.95/5.5	0.274	0.309	3.99	0.329	0.351	9.78	0.383	0.369	15.87
ASH	N 2.75/	0.344	0.357	1.21	0.420	0.391	3.33	0.447	0.403	5.05
	N 3/	0.350	0.356	1.77	0.420	0.393	5.13	0.450	0.401	7.37
	N 3.5/	0.350	0.353	1.81	0.421	0.396	5.82	0.453	0.398	7.65
	N 5/	0.348	0.353	4.01	0.421	0.401	12.86	0.451	0.398	16.94
	5 B 3/2	0.297	0.348	1.63	0.357	0.405	4.70	0.402	0.399	6.33
	5 B 4/4	0.258	0.340	2.51	0.313	0.400	7.27	0.358	0.399	9.24
	5 B 6/6	0.251	0.342	6.05	0.307	0.405	21.03	0.353	0.402	22.29
	5 G 3/2	0.323	0.383	1.42	0.401	0.431	4.07	0.423	0.424	5.65
	5 G 4/4	0.294	0.407	2.49	0.360	0.458	7.19	0.396	0.444	9.68
	5 G 4/6	0.283	0.427	3.01	0.345	0.478	8.88	0.381	0.458	11.01
	5 G 6/6	0.278	0.432	5.98	0.344	0.484	19.73	0.376	0.468	23.12
	5 Y 3/1	0.374	0.372	1.49	0.439	0.396	3.44	0.468	0.410	6.47
	5 Y 4/4	0.441	0.410	2.37	0.506	0.422	8.31	0.511	0.425	10.95
	5 Y 5/6	0.459	0.420	4.05	0.523	0.424	13.84	0.519	0.430	18.72
	5 Y 6/6	0.453	0.417	5.95	0.515	0.432	22.11	0.516	0.429	27.46
	3 YR 3.7/3.2	0.454	0.365	2.19	0.511	0.378	6.81	0.531	0.389	9.97
	5 R 2.5/2	0.420	0.346	1.02	0.480	0.371	2.94	0.510	0.381	4.52
	5 R 3/2	0.404	0.356	1.83	0.472	0.379	5.01	0.497	0.391	8.09
	5 R 3/4	0.474	0.339	1.52	0.535	0.358	5.51	0.550	0.367	6.87
	5 R 4/4	0.449	0.344	2.67	0.512	0.368	9.31	0.530	0.377	12.27
	5 R 4/6	0.490	0.336	2.59	0.548	0.357	9.71	0.563	0.363	12.28
	5 R 6/6	0.455	0.346	6.12	0.523	0.369	25.25	0.532	0.380	28.19
	5 P 3/2	0.352	0.328	1.69	0.423	0.369	4.39	0.455	0.379	6.94
	5 P 4/6	0.356	0.295	2.82	0.420	0.341	9.41	0.463	0.351	11.51
	5 P 6/6	0.351	0.311	6.92	0.406	0.348	23.67	0.457	0.364	28.68
	4.3 PB 4.95/5.5	0.271	0.308	3.95	0.340	0.354	11.88	0.376	0.376	15.05

(Table continues on next page)

Table 2. Continued

Observer	Test Panel	Test x	Test y	Test Y	Match x	Match y	Match Y	Const. x	Const. y	Const. Y
PBE	N 2.75/	0.370	0.350	0.98	0.429	0.382	2.45	0.456	0.398	4.30
	N 3/	0.374	0.359	1.41	0.428	0.393	3.56	0.458	0.404	6.22
	N 3.5/	0.374	0.354	1.56	0.420	0.390	3.88	0.458	0.399	6.71
	N 5/	0.373	0.355	3.65	0.423	0.393	8.11	0.459	0.400	16.51
	5 B 3/2	0.321	0.354	1.35	0.383	0.399	3.56	0.412	0.403	5.49
	5 B 4/4	0.274	0.345	2.15	0.334	0.386	4.79	0.366	0.404	8.71
	5 B 4/6	0.236	0.331	2.19	0.306	0.381	5.95	0.324	0.389	7.92
	5 B 6/6	0.269	0.347	5.23	0.303	0.384	11.50	0.359	0.405	20.30
	5 G 3/2	0.349	0.383	1.24	0.410	0.415	3.25	0.435	0.423	5.47
	5 G 4/4	0.317	0.407	2.12	0.381	0.432	5.21	0.404	0.444	8.68
	5 G 4/6	0.306	0.426	2.52	0.368	0.455	6.10	0.393	0.458	10.10
	5 G 6/6	0.299	0.422	4.73	0.350	0.458	11.38	0.387	0.460	19.60
	5 Y 3/1	0.403	0.373	1.22	0.449	0.397	2.91	0.479	0.410	5.67
	5 Y 4/4	0.466	0.405	2.02	0.490	0.415	4.51	0.517	0.425	9.65
	5 Y 5/6	0.483	0.415	3.52	0.511	0.422	7.24	0.525	0.429	17.51
	5 Y 6/6	0.476	0.411	5.09	0.492	0.418	13.50	0.520	0.428	23.84
	2.2 YR 6.47/4.12	0.457	0.362	6.73	0.492	0.393	17.90	0.523	0.389	30.13
	3 YR 3.7/3.2	0.479	0.359	1.91	0.510	0.383	4.57	0.537	0.387	8.88
	5 R 2.5/2	0.444	0.347	0.96	0.468	0.372	2.62	0.515	0.384	4.51
	5 R 3/2	0.435	0.351	1.57	0.493	0.377	3.91	0.508	0.387	7.12
	5 R 3/4	0.497	0.336	1.41	0.515	0.362	3.67	0.555	0.365	6.36
	5 R 4/4	0.472	0.343	2.36	0.498	0.366	5.42	0.535	0.376	10.88
	5 R 4/6	0.516	0.335	2.49	0.545	0.356	6.27	0.568	0.363	12.13
	5 R 6/6	0.476	0.344	5.43	0.514	0.371	12.72	0.537	0.378	25.78
	5 P 6/6	0.376	0.316	5.93	0.409	0.345	14.63	0.466	0.368	26.39
	4.3 PB 4.95/5.5	0.290	0.312	3.41	0.322	0.337	6.89	0.384	0.377	13.70
	5 PB 4/10	0.240	0.268	2.63	0.264	0.292	5.24	0.328	0.333	9.50
	5 PB 6/10	0.266	0.290	5.16	0.276	0.312	9.87	0.360	0.355	19.92
	5 YR 7/12	0.562	0.375	8.11	0.587	0.381	17.71	0.584	0.389	40.21
	5 YR 6/10	0.550	0.376	5.61	0.565	0.382	13.68	0.576	0.392	27.46
	5 YR 5/8	0.534	0.375	3.97	0.552	0.383	9.99	0.567	0.394	20.27
	5 G 5/8	0.282	0.446	3.82	0.322	0.487	7.68	0.368	0.476	14.71
	5 G 7/8	0.296	0.437	7.41	0.340	0.476	18.11	0.384	0.471	31.29
	5 Y 6/8	0.502	0.424	5.02	0.519	0.421	12.14	0.533	0.434	24.04

[a] The table provides the Munsell designations for the test surfaces, chromaticities and luminances of the test surfaces under the test illuminant, and chromaticities and luminances of the observers' asymmetric matches. Data are collapsed across match surfaces. The final three columns of the table provide a color constancy prediction based on the measurements and calculation described in the text (see Subsection 4.F). Chromaticities are 1931 CIE x and y. Luminances are cd/m^2.

One intriguing possibility is that our color experience at a location is described by more than three variables. This is possible if the influence of the illuminant (or, more generally, of the viewing context) has the effect of changing the perceptual representation of color in a way that cannot be compensated for simply by varying the tristimulus coordinates at a single location. Such an effect might be expected if the visual system uses color to code both surface and illuminant identity.[58]

Although our observation suggests that asymmetric matching cannot be used to characterize completely the effect of context on color appearance, we proceed nonetheless with an analysis of the matching data.

2. *Color Constancy—Comparison with Physical Measurements*

Were our observers color constant? For a color constant observer, the appearance of any surface should remain invariant as the surface is moved from one location in a room to another, independent of any illumination gradients in the room. We can use our asymmetric matching data to ask whether this was the case for our observers. Our asymmetric matches establish pairs of stimuli that appear identical when each is viewed in the appropriate location. If our observers were color constant, the match to any test surface should have the same tristimulus coordinates as the test surface would have when it was placed at the match location (with the colorimeter turned off). Since our experiment was conducted with real surfaces, we can test this color constancy prediction. For Experiment 2a we determined the color constancy prediction by moving our test surfaces to the match location and measuring their tristimulus coordinates under the match illuminant. For Experiment 2b we did not physically move the test surfaces to the match location. Instead, we used the full spectral measurements of test surfaces along with the full spectral measurements of the test and match illuminants to generate the color constancy prediction.

In particular, we calculated the spectrum $C_m(\lambda)$ that would have reached the observer if the test surface had been placed at the match location as $C_m(\lambda) = C_t(\lambda)E_m(\lambda)/E_t(\lambda)$, where $C_t(\lambda)$ is the spectrum of the light reaching the observer from the test surface when it is placed at the test location, $E_t(\lambda)$ is the spectrum of the test illuminant, and $E_m(\lambda)$ is the spectrum of the match illuminant. We then obtained tristimulus coordinates from $C_m(\lambda)$.

The constancy predictions for all of our test surfaces are provided along with the matching data in Tables 1 and 2. Figure 7 compares WAB's asymmetric matches to these physical measurements for a few test surfaces. The points shown are typical of those for all test surfaces. We see that the matching coordinates lie between the coordinates of the test surface under the two illuminants. Our results represent neither complete constancy nor a complete absence of constancy. Rather, the visual system compensates partially for the illumination gradient. In Subsection 6.A we analyze quantitatively the degree to which our observers were color constant.

3. Control of Adaptation

In our experiment the observers were free to look back and forth across the illumination gradient. This means that the state of local chromatic adaptation was not under experimental control. The reason for this design was to allow our experiments to capture performance that would occur in a natural setting, where observers are free to scan their environment. We believe that this is sensible design for the purpose of characterizing natural performance. On the other hand, our experiments do not allow us to isolate the role of any particular adaptive mechanism, such as local chromatic adaptation, which might mediate the performance we observe.

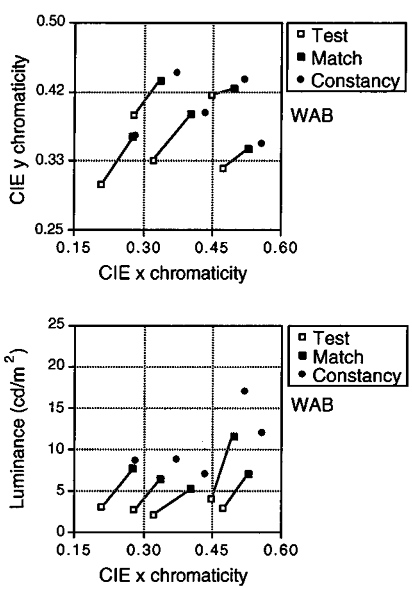

Fig. 7. Color constancy, Experiment 2a: comparison of WAB's asymmetric matches from Experiment 2a with physical measurements of the light reflected from the test surface when the surface is viewed under the match illuminant. Open squares, test surface coordinates; solid squares, match surface coordinates; solid circles, coordinates of the test surfaces under the match illuminant. If the observer were color constant, the solid squares and the solid circles would coincide. To avoid overwhelming the plot, results are shown only for a few typical test surfaces.

5. MODELING THE EFFECT OF THE ILLUMINATION CHANGE

Our observers were not completely color constant, nor were their matches colorimetric matches. Thus it becomes interesting to develop a model that predicts their performance. We consider four models. Let the three-dimensional column vectors $\mathbf{r}_t$ and $\mathbf{r}_m$ represent the cone excitation coordinates of the light reaching the observer from the test and match surfaces. Our most general model is the *affine model*. For the affine model, we write

$$\mathbf{r}_m = \mathbf{A}\mathbf{r}_t + \mathbf{a}, \tag{1}$$

where $\mathbf{A}$ is a 3×3 matrix and $\mathbf{a}$ is a three-dimensional column vector. In the affine model the entries of the matrix $\mathbf{A}$ are parameters that can describe multiplicative gain changes both at the receptors and after an opponent transformation. The entries of the vector $\mathbf{a}$ are parameters that describe an additive process. The affine model can be thought of as an instantiation of the two-process model of Jameson and Hurvich.[11] It was used by Burnham *et al.*[10] to describe an extensive set of asymmetric matches collected haploscopically.

Our *linear model* is a special case of the affine model where the vector $\mathbf{a}$ is constrained to be zero, so that

$$\mathbf{r}_m = \mathbf{A}\mathbf{r}_t\,. \tag{2}$$

Our *diagonal model* specializes still further by requiring that the matrix $\mathbf{A}$ be diagonal. The diagonal model allows for multiplicative gain changes that are specific to the three classes of cones but does not incorporate any independent gain control after an opponent transformation.

The affine, linear, and diagonal models are motivated by thinking about visual mechanisms. Our final model is motivated by a more computational approach. We take seriously the notion that the visual system is trying to achieve constancy, but we allow for the possibility that it does not achieve it perfectly. Consider a series of estimation steps sufficient for color constant performance in our asymmetric matching task. The first step is to estimate the test illuminant. Given this estimate, the second step is to estimate the reflectance function of the test surface from the cone excitation coordinates of the light reflected from it. For color constant performance, this surface must have the same appearance when it is viewed under the match illuminant. Thus the third step is to estimate the match illuminant and compute the light that would be reflected from the test surface under it. If all three of these estimations are performed accurately, an observer can achieve color constancy by adjusting the match surface to have the same cone coordinates as the light computed in the third step. The estimations are known to be computationally difficult, however.[47,48] Even if the visual system is attempting this type of computation, it may make errors.

Suppose that the visual system correctly performs the first estimation step, so that it has access to an accurate representation of the test illuminant. By using techniques based on linear model representations of physical spectra, it is possible for the visual system to reconstruct a reasonable estimate of the test surface reflectance.[46] Now suppose that the visual system makes an error in estimating the match illuminant and then uses this estimate to determine the asymmetric match. Although the resulting match will represent color constant performance with respect to the estimated illuminant, it will not represent constancy with respect to the actual match illuminant. Our *equivalent illuminant model* is based on this idea. We assume that the only error the visual system makes in an attempt to achieve constancy is that it misestimates the match illuminant. We call the visual system's estimate the equivalent illuminant. For any given equivalent illuminant, the model predicts the asymmetric match for any test surface. In Appendix A we make the equivalent illuminant model explicit and show that it is a special case of the linear model.

A. Fitting the Models

The affine, linear, and diagonal models can be fitted to the data analytically by using standard regression procedures. These procedures minimize the squared error of the fit when the error is computed in cone excitation coordinates. It is well known that cone excitation space is not perceptually uniform. For this reason, we fitted our models by using numerical search procedures. The procedures determined the model parameter values that minimized the mean CIELAB ΔE^* color difference between the predictions and the data, where the average was taken across each asymmetric matching data set. We used the routines in the MATLAB Optimization Toolbox[59,60] to perform the minimization. Our procedure for fitting the equivalent illuminant model minimizes the same error measure and is described in Appendix A.

B. Model Fits

The histogram in Fig. 8(a) summarizes how well each of our models fits the data of Experiment 2a, and the one in Fig. 8(b) summarizes the fits for Experiment 2b. The histograms show the mean CIELAB ΔE^* prediction error obtained with each of our four models. To compute the errors, we predicted the asymmetric match for each test surface, computed the error to each individual match, and averaged over the data set. This is the same measure that was minimized in our model fitting procedures. Along with our model fits, the histograms show comparison measures that indicate the precision of the observers' matches (determined from replicated matches, labeled "precision"), the quality of fit that would be obtained if one predicted the asymmetric matches with the cone coordinates of the test surface (labeled "no effect"), and the quality of fit that would be obtained by assuming the observers were perfectly color constant (labeled "constancy").

Several conclusions can be drawn from Fig. 8. First, all four models describe a large fraction of the asymmetric matching effect: the prediction errors are much smaller than those that would be obtained by assuming no effect. Second, the fits of all four models are quite similar. If we average across the five observers, the CIELAB ΔE^* predictive errors for the affine, linear, diagonal, and equivalent illuminant models are 3.41, 3.58, 4.21, and 4.23, respectively. A just-noticeable difference under optimal conditions is approximately 1 ΔE^* unit, and the mean precision of our observers' asymmetric matches was 1.91 ΔE^* units. The third conclusion that may be drawn from the figure is that none of our models describes all of the regularity in the data.

Figure 9 provides a more detailed look at the model fits. The figures collapse the data for all five observers in Experiments 2a and 2b. Each subplot in the figure shows a scatterplot of predicted match cone coordinates versus observed match cone coordinates. The three columns show data for the L, M, and S cones respectively. The top row shows the case in which each individual match is predicted by the mean match for its test surface. The scatter around the diagonal is a visual indication of the precision of the data. The next four rows show the plots for our affine, linear, diagonal, and equivalent illuminant models. If the models fitted to the precision of the data, the plotted points would cluster around the diagonal to the same extent as in the top row. We see that the four models provide essentially the same quality fit and also that this fit is only slightly worse than the precision of the data.

For all four models, the points in Fig. 9 tend to lie below the diagonal for high observed match values. Recall, however, that we fitted our models by minimizing error in the CIELAB ΔE^* uniform color space. We refitted the linear and diagonal models to minimize squared error in cone excitation space. Figure 10 shows the corresponding scatterplots. In these plots the trend is greatly reduced. Thus we believe that the trends seen in Fig. 9 do not indicate systematic model failure.

The linear model is a special case of the affine model, so its predictive error must be larger than that of the affine model. The linear model has 9 free parameters, while the affine model has 12. The average difference in predictive error between these two models is very small (0.18 ΔE^* unit). This does not mean that there is no neural addition to the signals transmitted by cones but rather that our experimental manipulation did not substantially change any such addition.

The diagonal and equivalent illuminant models are both special cases of the linear model, and their predictive errors must be larger than that of the linear model. Both the diagonal and equivalent illuminant models have three free parameters, while the linear model has nine. The average fit of these models is quite good in absolute terms and not much worse than that of the linear model (0.62 and 0.64 ΔE^* unit, respectively). For descriptive and predictive purposes, we currently favor the diagonal and equivalent illuminant models because of their small number of parameters. In principle, the parametric values for either of these models may be established with a single asymmetric match. In practice, a small number matches would probably be sufficient.

To provide a sense of the quality of the fit obtained with the diagonal and equivalent illuminant models, in Fig. 11 we compare WAB's asymmetric matches from Experiment 2a with the predictions of these models for a few test surfaces. To avoid graph clutter, only a few points are shown, but the points shown are typical. The predic-

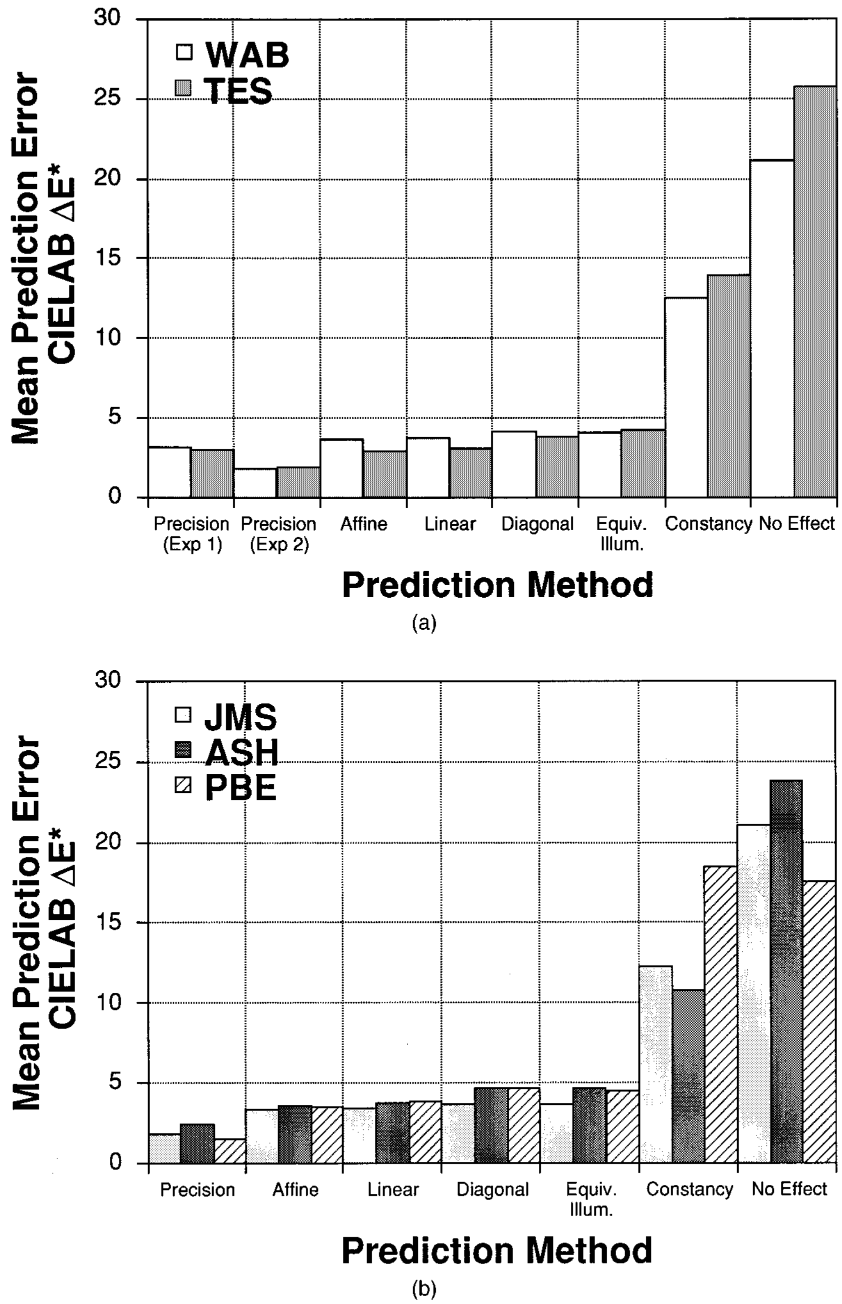

Fig. 8. Experiment 2 model fit. Each bar in the histograms gives the mean CIELAB ΔE^* prediction error for one model and observer. (a) Fits for Experiment 2a for observers WAB and TES; (b) fit for Experiment 2b for observers JMS, ASH, and PBE. Precision: each individual match is fitted with the mean match for the corresponding test surface, giving a bound on how well any model can fit the data. In (a) we show the precision both for Experiment 2a and for the symmetric matching conditions of Experiment 1. Affine: each individual match is predicted by the affine model applied to the cone coordinates of the corresponding test surface. Linear: each individual match is predicted by the linear model. Diagonal: diagonal model predictions. Equiv. Illum.: equivalent illuminant model predictions. Constancy: each individual match is predicted by the measurement of the corresponding color constancy prediction (see Subsection 4.F). No effect: each individual match is predicted by the cone excitation coordinates of the corresponding test surface, giving an estimate of the size of the effect of changing the illuminant. The CIELAB coordinates of the matches and predictions were computed with respect to a white point defined by the illuminant at the match location. Note that this illuminant differed between Experiments 1 and 2. Direct comparisons of CIELAB ΔE^* values across different white points are less meaningful than comparisons of values computed for the same white point, since the accuracy of across-white-point comparisons depends critically on the adaptation model incorporated in the CIELAB calculation. The apparently lower match precision in Experiment 1 in comparison with that of Experiment 2 is probably not significant. The key comparisons in the figure are all of CIELAB ΔE^* values computed for a single white point.

tions of the two models are very similar and lie close to the data.

The small difference between the diagonal and the equivalent illuminant models (0.02 ΔE^* unit) does not lead us to prefer one over the other. At present, choosing between the diagonal and the equivalent illuminant models is primarily a matter of taste. The diagonal model has a natural interpretation in terms of visual mecha-

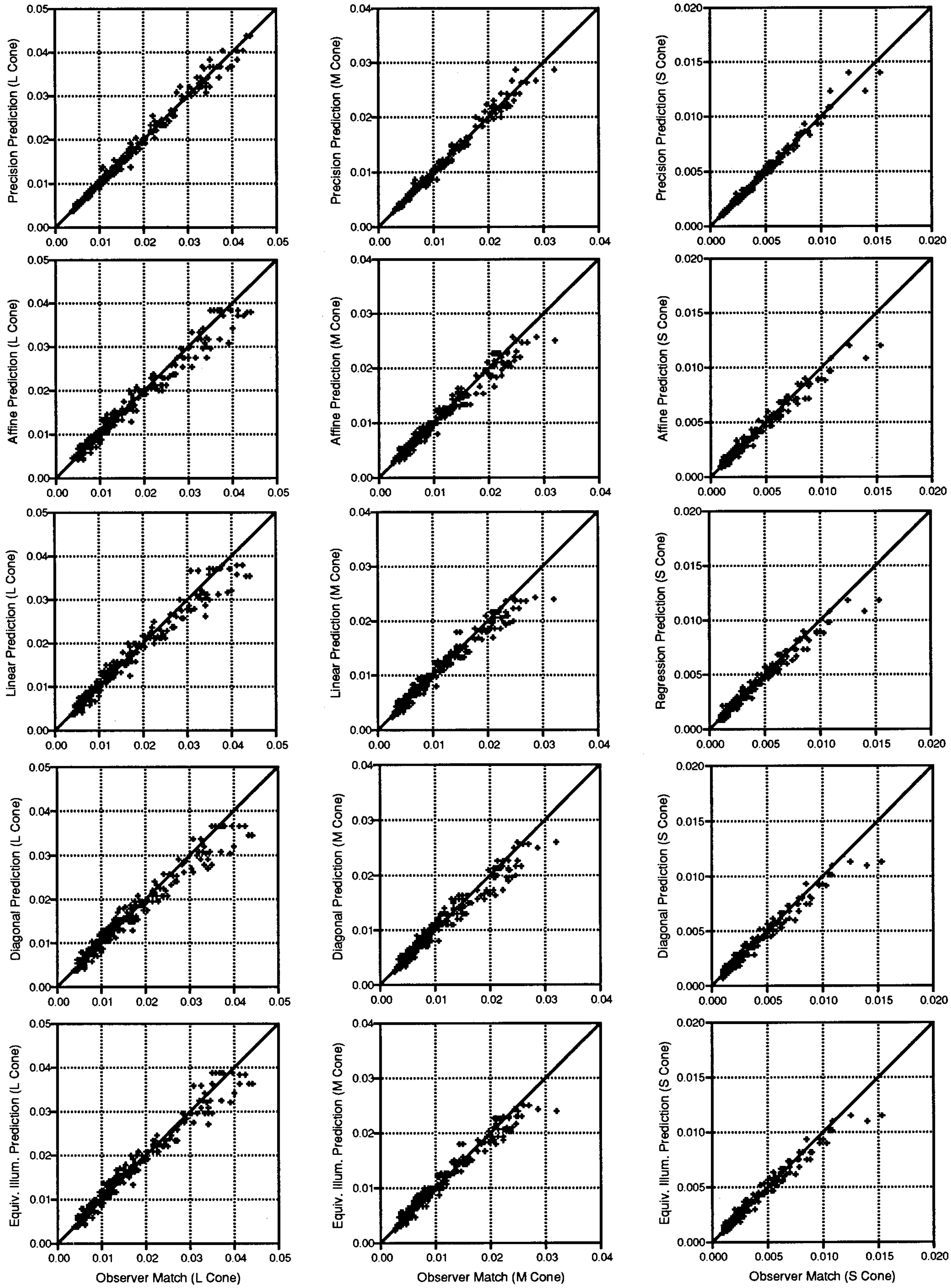

Fig. 9. Scatterplots of model fits. Each subplot shows a scatterplot of predicted versus observed match cone coordinates. Each column shows data for L, M, or S cones. Top row, each individual match is predicted by the mean match for its test surface; next four rows, affine, linear, diagonal, and equivalent illuminant models. The model fitting procedure minimized CIELAB ΔE* predictive error. Data are for all five observers in Experiments 2a and 2b. Each observer's data was fitted separately.

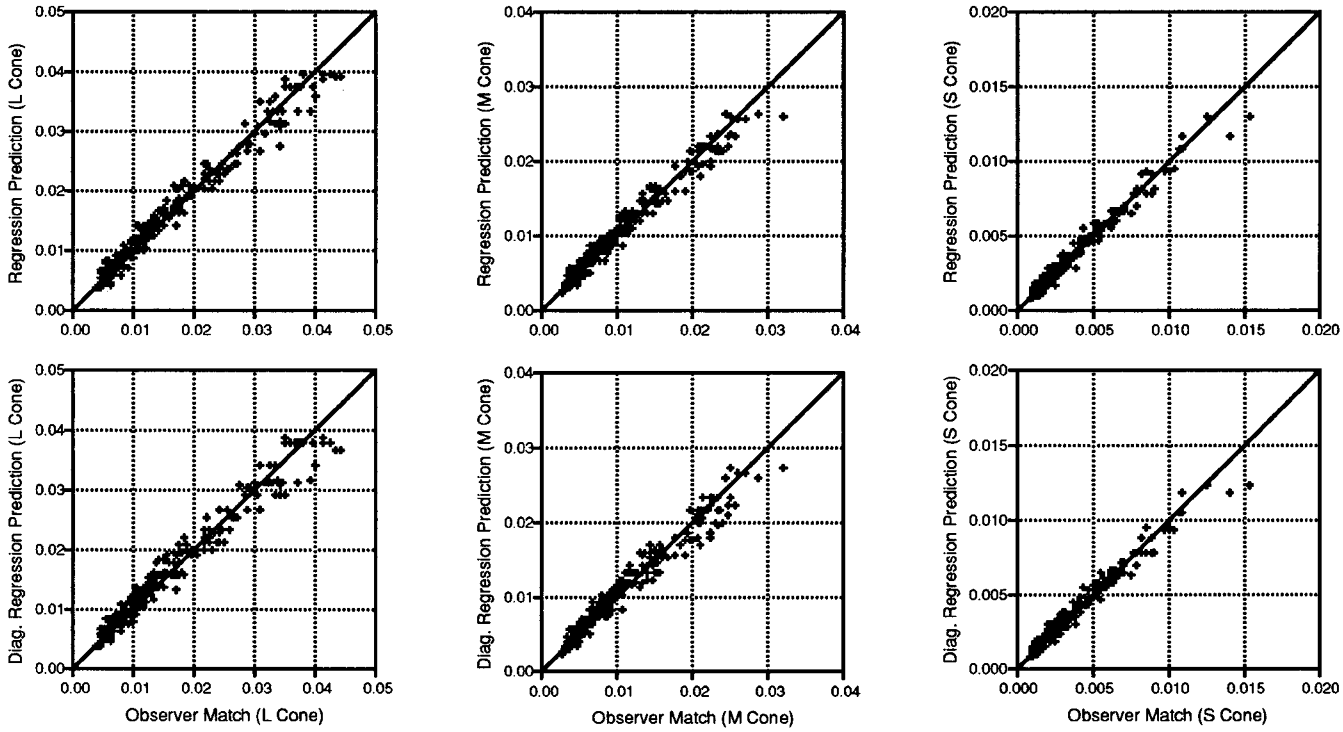

Fig. 10. Scatterplots of model fits. Each subplot shows a scatterplot of predicted versus observed match cone coordinates. Each column shows data for L, M, or S cones. Top row, linear model; bottom row, diagonal model. The model fitting procedure minimized the squared predictive error in cone excitation space. Data are for all five observers in Experiments 2a and 2b. Each observer's data was fitted separately.

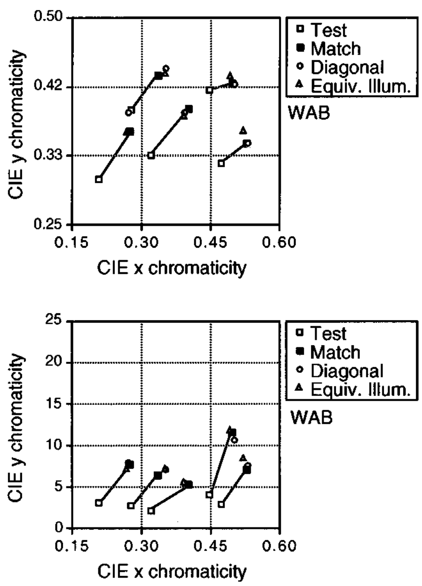

Fig. 11. Model fit: comparison of WAB's asymmetric matches from Experiment 2a with the predictions of the diagonal and equivalent illuminant models. Open squares, test surface coordinates; solid squares, match surface coordinates; open circles, diagonal model; open triangles, equivalent illuminant model. To avoid overwhelming the plot, results are shown only for a few typical test surfaces.

nisms (namely, that gain control occurs in cone-specific pathways), while the equivalent illuminant model is more closely linked to computational models of color constancy (see below).

6. GENERAL DISCUSSION

A. Degree of Color Constancy

Our equivalent illuminant model summarizes the entire set of asymmetric matching data by specifying the equivalent illuminant. Figure 12 shows the equivalent illuminants derived from our data for three of our observers in Experiments 2a and 2b. For comparison, the actual test and match illuminants are also shown. Table 3 provides the equivalent illuminants for all of our observers, along with the corresponding test and match illuminants.

Since the equivalent illuminant summarizes the data, we can use it to compute a color constancy index

$$\mathrm{CI} = 1 - \frac{|\mathbf{e}_m - \hat{\mathbf{e}}_{em}|}{|\mathbf{e}_m - \mathbf{e}_t|}, \tag{3}$$

where $\mathbf{e}_m$ is a vector representing the match illuminant, $\mathbf{e}_t$ is a vector representing the test illuminant, and $\hat{\mathbf{e}}_{em}$ is a vector representing the equivalent illuminant. This index compares the separation between the match illuminant and the equivalent illuminant with the separation between the match illuminant and the test illuminant. To perform calculations, we specified the illuminants in the CIELAB L*a*b* uniform color space (relative to a white point defined by the test illuminant).

The constancy indices obtained in Experiments 2a and 2b are given in Table 3. The mean constancy index was 0.61. There is no difference between the mean value of the index for Experiments 2a and 2b, suggesting that the addition of extra Munsell panels in Experiment 2b had

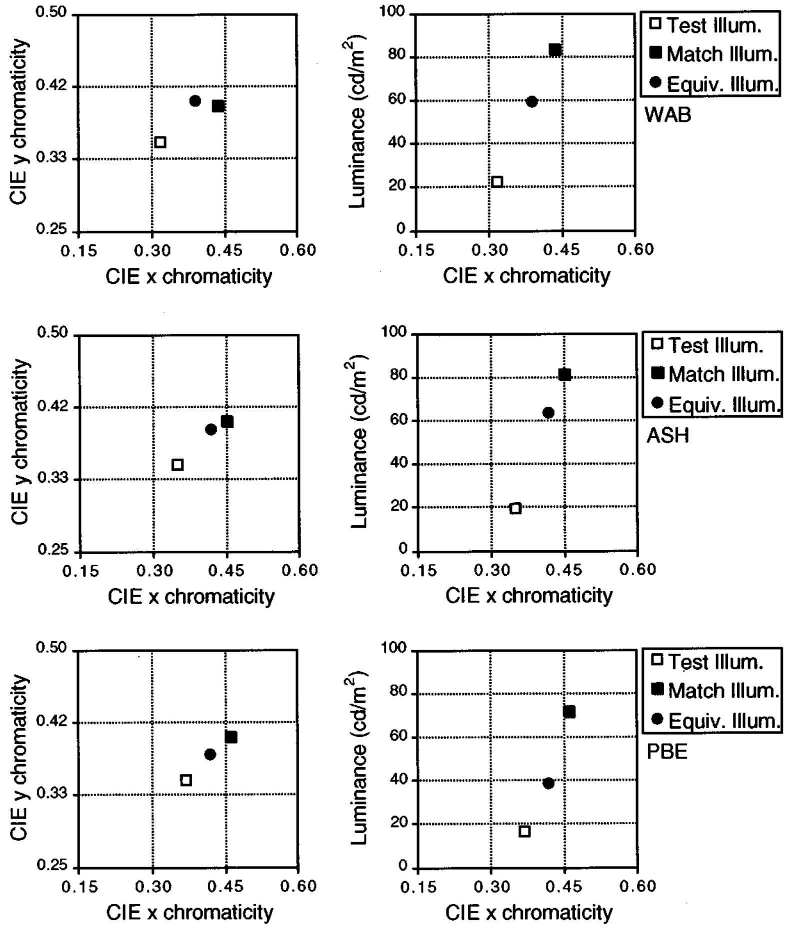

Fig. 12. Equivalent illuminants. Each row shows the coordinates of the equivalent illuminant (solid circles) for one observer. For comparison, the coordinates of the test illuminants (open squares) and match illuminants (solid squares) are also shown. If the observer were perfectly color constant, the solid circles would superpose on the closed squares. Top row, Experiment 2a, observer WAB; middle row, Experiment 2b, observer ASH; bottom row, Experiment 2b, observer PBE.

Table 3. Experimental Illuminants, Equivalent Illuminants, and Constancy Indices for Experiments 2a and b[a]

Illuminant	WAB	TES	JMS	ASH	PBE
Test x	0.318	0.317	0.346	0.351	0.371
Test y	0.352	0.353	0.367	0.354	0.353
Test Y	22.16	22.33	20.72	19.49	16.46
Match x	0.435	0.434	0.450	0.452	0.456
Match y	0.394	0.394	0.401	0.398	0.398
Match Y	83.23	83.32	88.36	81.60	71.96
Equiv. x	0.391	0.388	0.421	0.422	0.416
Equiv. y	0.399	0.393	0.404	0.393	0.384
Equiv. Y	59.29	69.69	57.61	63.78	38.86
Constancy index	0.59	0.63	0.64	0.70	0.48
Chromaticity only	0.57	0.57	0.68	0.69	0.52
Luminance only	0.60	0.78	0.55	0.71	0.40

[a] The top portion of the table provides the chromaticities and luminances of the measured test and match illuminants for each observer. It also gives the chromaticity and luminance of the equivalent illuminant derived from each observer's asymmetric matches. Chromaticities are 1931 CIE x and y. Luminances are cd/m^2. The bottom portion of the table gives the constancy indices computed for each observer. The overall constancy indices were computed by expressing the illuminants in the CIELAB L*a*b* uniform color space. The chromaticity indices were computed from the CIE $u'v'$ chromaticities of the illuminants. The luminance indices were computed from the luminances of the illuminants. The mean values of the constancy, chromaticity, and luminance indices are 0.61, 0.60, and 0.61, respectively.

little effect on the asymmetric matches. This may be because the viewing conditions for Experiment 2a were already quite rich.

If we separate the chromaticities of the illuminants from their luminances, we can compute chromaticity (CIE $u'v'$) and luminance constancy indices. Table 3 provides the value of these indices for each of our observers. The mean values of these were 0.60 and 0.61, respectively, so that our data show no quantitative difference in how well the visual system compensates for changes in illuminant luminance and how well it compensates for changes in illuminant chromaticity.

It is possible to compute a constancy index with our other models as well. To do so, we substitute for the equivalent illuminant in Eq. (3) a quantity obtained through applying the affine, linear, or diagonal mapping to the cone coordinates of the test illuminant. The resulting index values are essentially the same as those we obtained with the index that is based on the equivalent illuminant.

Our constancy index is related to the one introduced by Arend and his colleagues[6,27] (see also Ref. 7). Two differences are worth noting. First, our experiment is not restricted to the isoluminant plane, so we compute distances by using a three-dimensional uniform color space rather than a uniform chromaticity diagram. Second, we base our index on a model fit that summarizes the entire data set, rather than averaging indices computed individually for each test surface. Brainard and Wandell[3] reported constancy indices based on equivalent illuminant calculations. In their case, however, the equivalent illuminant was calculated independently for each test surface and the resulting indices were averaged over test surfaces.

B. Equivalent-illuminant Model

The equivalent-illuminant analysis summarizes the asymmetric matches with the visual system's estimate of the illuminant. The attractive feature of this approach is that it provides a link between human performance and computational models of color constancy, since the key to most of these models is how they estimate the illuminant.[46–48,61–71] Indeed, comparing the dependence of an observer's equivalent illuminant on stimulus conditions with the corresponding dependence of a model's illuminant estimates is a promising approach for testing whether computational models can describe human performance.

Our equivalent illuminant model is closely related to that of Speigle and Brainard,[72] who used essentially the same principles to predict when stimuli would appear self-luminous. Brainard and Wandell[3] also used equivalent illuminant ideas to model asymmetric color matches. They calculated a separate equivalent illuminant for each test surface, however, so that the value of the approach was to allow easy interpretation of individual matches in terms of color constancy rather than to provide a quantitative model for an entire matching data set. Arend[6] used a calculation closely related to that of Brainard and Wandell.[3]

Lucassen and Walraven[7] compare entire sets of asymmetric matches with the results of a calculation similar to our equivalent illuminant model. Their approach differs from ours in an important way. We treat the visual system's estimate of the illuminant as a parameter, so that we test the idea that the visual system estimates an illuminant without committing ourselves to how this estimate is obtained. They test a more specific computational model by using the gray-world assumption and the method of Buchsbaum[46] to obtain an illuminant estimate from their stimulus set.

C. Comparison with CRT-Based Experiments

One of the goals of our research is to understand the relation between performance under nearly natural conditions and performance measured by using CRT simulations of illuminated surfaces. Definite conclusions will require experiments that make direct comparisons of the two types of experiment with the use of methods, observers, and stimuli that are as closely matched as possible. Such experiments are under way in our laboratory.[73] Some comparisons can be made, however, between the present data and published data obtained with CRT simulations.

Brainard and Wandell[3,15] measured successive color constancy with an asymmetric matching method and then fitted versions of the affine, linear, and diagonal models to their data. Their conclusions match ours. All three models provided a good description of the data, and the fit of the reduced diagonal model was not substantially worse than the fit of the two more-general models (see Ref. 15, p. 1438). This suggests that empirical regularities found with CRT simulations can indeed be generalized to natural viewing conditions and that perhaps that the same mechanisms mediate performance in both situations. Brainard and Wandell also tested an illuminant linearity property that is not addressed by our data.[3,15]

Of the CRT-based experiments reported in the literature, the simultaneous asymmetric matching experiments of Arend and colleagues[26,27] are procedurally most similar to ours. In particular, we believe that our observers were performing their unasserted-color task rather than their perceived-surface-color task (see Section 2). Arend *et al.*[27] computed a color constancy index comparable to the one we used and reported this value for both of their tasks. We averaged their index values for their four observers and two experiments (see graph in Ref. 27, p. 667). For their unasserted-color task, the mean value of the index was 0.25.[74] The comparable value for our experiments is 0.60 (mean chromaticity index; see Table 3). The quantitative difference in constancy indices between our experiments and those of Arend and colleagues[26,27] suggests that there are important differences between the experiments that have been performed with CRT simulations and the present experiments with nearly natural scenes. These differences may arise because of differences in low-level attributes such as the size and luminance of the stimuli, because of the increased richness of the natural scenes relative to the simple CRT simulations, or because the visual system treats real and rendered images differently. Discovering exactly what factors are most important is a high priority for future research.

APPENDIX A: EQUIVALENT ILLUMINANT MODEL

In this appendix we describe the equivalent illuminant model. We use the general linear-model-based colorimetric techniques described by Brainard.[53]

1. Preliminaries

Let $\mathbf{s}$ be an N_λ-dimensional column vector whose entries specify a surface-reflectance function sampled at N_λ evenly spaced wavelengths throughout the visible spectrum. The vector $\mathbf{s}$ lies within an N_s-dimensional linear model if we can write

$$\mathbf{s} = \mathbf{B}_s \mathbf{w}_s , \qquad \text{(A1)}$$

where $\mathbf{B}_s$ is a fixed $N_\lambda \times N_s$-dimensional matrix. We call $\mathbf{B}_s$ the basis matrix of the linear model; we call the columns of $\mathbf{B}_s$ the basis vectors of the linear model. A linear model is specified completely by its basis matrix. Once a basis matrix for surface spectra has been chosen, we can specify any vector within the linear model by its vector of weights $\mathbf{w}_s$. In parallel fashion, let $\mathbf{e}$ be an N_λ-dimensional column vector whose entries specify illuminant power sampled at N_λ evenly spaced wavelengths throughout the visible spectrum. The vector $\mathbf{e}$ lies within an N_e-dimensional linear model if we can write

$$\mathbf{e} = \mathbf{B}_e \mathbf{w}_e , \qquad \text{(A2)}$$

where $\mathbf{B}_e$ is a fixed $N_\lambda \times N_e$-dimensional matrix. In our calculations, we assume that both surface and illuminant spectra are constrained to lie within three-dimensional linear models and that the basis matrices for the two linear models are known.

Let the $3 \times N_\lambda$-dimensional matrix $\mathbf{T}$ be the matrix whose rows contain the L-, M-, and S-cone spectral sensitivities sampled at evenly spaced wavelengths throughout the visible spectrum. In this paper we use the Smith–Pokorny estimates of the cone spectral sensitivities.[75,76] The cone excitation coordinates $\mathbf{r}$ of the light reflected from a surface $\mathbf{w}_s$ under an illuminant $\mathbf{w}_e$ are given by

$$\mathbf{r} = \mathbf{T}\ \mathrm{diag}(\mathbf{B}_e \mathbf{w}_e) \mathbf{B}_s \mathbf{w}_s , \qquad \text{(A3)}$$

where the expression diag $(\mathbf{B}_e \mathbf{w}_e)$ indicates the diagonal matrix with the entries of the vector $\mathbf{B}_e \mathbf{w}_e$ along its main diagonal.

2. Equivalent Illuminant Model

Suppose that we choose three-dimensional linear models $\mathbf{B}_s$ and $\mathbf{B}_e$ for illuminants and surfaces, and suppose that the observer estimates the test illuminant within $\mathbf{B}_e$ to be $\hat{\mathbf{w}}_{et}$. If the tristimulus coordinates of the light reflected from the test surface are $\mathbf{r}_t$, then the visual system can estimate the test surface reflectance $\hat{\mathbf{w}}_s$ within the linear model $\mathbf{B}_s$ by inverting Eq. (A3):

$$\hat{\mathbf{w}}_s = [\mathbf{T}\ \mathrm{diag}(\mathbf{B}_e \hat{\mathbf{w}}_{et}) \mathbf{B}_s]^{-1} \mathbf{r}_t . \qquad \text{(A4)}$$

This inversion method for estimating surface reflectance is due to Buchsbaum.[46]

Now suppose that the observer estimates the match illuminant within $\mathbf{B}_e$ to be $\hat{\mathbf{w}}_{em}$. Given these illuminant estimates, the observer can achieve color constancy if the surface $\hat{\mathbf{w}}_s$ rendered under the illuminant $\hat{\mathbf{w}}_{em}$ has the same appearance as the surface $\hat{\mathbf{w}}_s$ rendered under $\hat{\mathbf{w}}_{et}$. This condition implies that in our asymmetric matching experiment the light reflected from the match surface should have cone excitation coordinates

$$\begin{aligned}\mathbf{r}_m &= \mathbf{T}\ \mathrm{diag}(\mathbf{B}_e \hat{\mathbf{w}}_{em}) \mathbf{B}_s \hat{\mathbf{w}}_s \\ &= \mathbf{T}\ \mathrm{diag}(\mathbf{B}_e \hat{\mathbf{w}}_{em}) \mathbf{B}_s [\mathbf{T}\ \mathrm{diag}(\mathbf{B}_e \hat{\mathbf{w}}_{et}) \mathbf{B}_s]^{-1} \mathbf{r}_t \\ &= \mathbf{A}(\hat{\mathbf{w}}_{et}, \hat{\mathbf{w}}_{em}) \mathbf{r}_t , \end{aligned} \qquad \text{(A5)}$$

where $\mathbf{A}(\hat{\mathbf{w}}_{et}, \hat{\mathbf{w}}_{em})$ is a 3×3 matrix that depends on the illuminant estimates $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$ but not on the test surface.

The first implication of Eq. (A5) is that once we choose illuminant estimates $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$, the model predicts the asymmetric match for any test surface. The second implication of Eq. (A5) is that, like the diagonal model, the equivalent illuminant model is a special case of the linear model. In the diagonal model the general matrix $\mathbf{A}$ is constrained to be diagonal, and thus it has three degrees of freedom. In the equivalent illuminant model the matrix $\mathbf{A}$ is constrained by Eq. (A5). When we hold $\hat{\mathbf{w}}_{et}$ fixed, the equivalent illuminant matrix $\mathbf{A}(\hat{\mathbf{w}}_{et}, \hat{\mathbf{w}}_{em})$ also has three degrees of freedom.

We can fit the equivalent illuminant model to data by treating $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$ as parameters. These parameters are not independently identifiable, however. If we multiply both $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$ by the same scale factor, the predicted mapping $\mathbf{A}(\hat{\mathbf{w}}_{et}, \hat{\mathbf{w}}_{em})$ does not change. In practice, we have found that the data are not powerful enough to identify reliably both $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$ even if we fix the free scale factor: the model predictions depend jointly on $\hat{\mathbf{w}}_{et}$ and $\hat{\mathbf{w}}_{em}$ in such a way that we can trade off changes in the two without much effect on the predictions. It is not clear to us what properties of the model or of our data sets lead to this lack of identifiability. To fit the model to our data, however, we fix $\hat{\mathbf{w}}_{et}$ as described below and vary only $\hat{\mathbf{w}}_{em}$ in our parameter search.

A side condition of the equivalent illuminant model is the choice of linear models $\mathbf{B}_s$ and $\mathbf{B}_e$. Indeed, if these matrices are not specified *a priori*, the equivalent illuminant model has little predictive power. In our calculations we chose $\mathbf{B}_s$ to be a three-dimensional linear model for the Munsell papers computed from the spectral measurements of Kelly *et al.*,[77] Nickerson,[78] and Nickerson and Wilson.[79] We chose $\mathbf{B}_e$ as a three-dimensional linear model that approximated the actual illuminants in our experimental room.

To find the equivalent illuminant that provided the best fit to the data, we implemented a numerical search procedure over the three-dimensional space of equivalent illuminants $\hat{\mathbf{w}}_{em}$. For this search we fixed $\hat{\mathbf{w}}_{et}$ so that it was the best approximation to the actual test illuminant within the linear model $\mathbf{B}_e$. For each choice of $\hat{\mathbf{w}}_{em}$ we computed the predicted matches through Eq. (A5) and used the routines in the MATLAB Optimization Toolbox[59,60] to find the value of $\hat{\mathbf{w}}_{em}$ that minimized the prediction error. Although we express the model in terms of cone excitation coordinates, we used the CIELAB ΔE^* color difference metric to define the prediction error. We chose the model parameters that minimized the mean value of this error across each asymmetric matching data set.

ACKNOWLEDGMENTS

J. D. Tietz provided technical support. We thank A. A. Berkman (pilot experiments), P. B. Egan, A. S. Haque, and T. E. Sitzer for careful observing. K. H. Bauml, P. B. Egan, J. M. Kraft, and two anonymous reviewers provided critical comments on the manuscript. Experiments 1 and 2a are described in W. A. Brunt's Master's thesis. Experiment 2b was conducted as part of J. M. Speigle's Ph.D. thesis research and is described briefly by Speigle and Brainard.[80] Preliminary versions of this research were presented at the 1995 Trieste Encounter in Cognitive Science and the 1996 Association for Research in Vision and Ophthalmology Meeting.[81,82] The work was supported by National Eye Institute grant EY 10016.

REFERENCES AND NOTES

1. E. G. Boring, *Sensation and Perception in the History of Experimental Psychology* (D. Appleton Century, New York, 1942).
2. R. M. Evans, *An Introduction to Color* (Wiley, New York, 1948).
3. D. H. Brainard and B. A. Wandell, "A bilinear model of the illuminant's effect on color appearance," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon eds. (MIT Press, Cambridge, Mass., 1991).
4. D. H. Brainard, B. A. Wandell, and E.-J. Chichilnisky, "Color constancy: from physics to appearance," Curr. Direc. Psychol. Sci. **2**, 165–170 (1993).
5. M. D. Fairchild and P. Lennie, "Chromatic adaptation to natural and incandescent illuminants," Vision Res. **32**, 2077–2085 (1992).
6. L. E. Arend, Jr., "How much does illuminant color affect unattributed colors?" J. Opt. Soc. Am. A **10**, 2134–2147 (1993).
7. M. P. Lucassen and J. Walraven, "Color constancy under natural and artificial illumination," Vision Res. **36**, 2699–2711 (1996).
8. J. von Kries, "Chromatic adaptation," in *Sources of Color Vision*, D. L. MacAdam ed. (MIT Press, Cambridge, Mass., 1970). Originally published in *Festschrift der Albrecht-Ludwigs-Universitat* (1902), pp. 145–148.
9. D. B. Judd, "Hue saturation and lightness of surface colors with chromatic illumination," J. Opt. Soc. Am. **30**, 2–32 (1940).
10. R. W. Burnham, R. M. Evans, and S. M. Newhall, "Prediction of color appearance with different adaptation illuminations," J. Opt. Soc. Am. **47**, 35–42 (1957).
11. D. Jameson and L. M. Hurvich, "Theory of brightness and color contrast in human vision," Vision Res. **4**, 135–154 (1964).
12. J. J. McCann, S. P. McKee, and T. H. Taylor, "Quantitative studies in retinex theory: a comparison between theoretical predictions and observer responses to the 'color Mondrian' experiments," Vision Res. **16**, 445–458 (1976).
13. J. Walraven, "Discounting the background: the missing link in the explanation of chromatic induction," Vision Res. **16**, 289–295 (1976).
14. S. K. Shevell, "The dual role of chromatic backgrounds in color perception," Vision Res. **18**, 1649–1661 (1978).
15. D. H. Brainard and B. A. Wandell, "Asymmetric color-matching: how color appearance depends on the illuminant," J. Opt. Soc. Am. A **9**, 1433–1448 (1992).
16. K. H. Bauml, "Color appearance: effects of illumant changes under different surface collections," J. Opt. Soc. Am. A **11**, 531–542 (1994).
17. K. H. Bauml, "Illuminant changes under different surface collections: examining some principles of color appearance," J. Opt. Soc. Am. A **12**, 261–271 (1995).
18. E. J. Chichilnisky and B. A. Wandell, "Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching," Vision Res. **35**, 239–254 (1995).
19. E. J. Chichilnisky and B. A. Wandell, "Seeing gray through the on and off pathways," Visual Neurosi. **13**, 591–596 (1996).
20. Y. Nayatani, K. Takahama, H. Sobagaki, and K. Hashimoto, "Color-appearance model and chromatic adaptation transform," Color Res. Appl. **15**, 210–221 (1990).
21. R. W. G. Hunt, "Revised colour-appearance model for related and unrelated colours," Color Res. Appl. **16**, 146–165 (1991).
22. M. D. Fairchild and R. S. Berns, "Image color-appearance specification through extension of CIELAB," Color Res. Appl. **18**, 178–190 (1993).
23. R. W. G. Hunt, "Light and dark adaptation and the perception of color," J. Opt. Soc. Am. **42**, 190–199 (1952).
24. L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," Psychol. Rev. **64**, 384–404 (1957).
25. A. B. Poirson and B. A. Wandell, "Appearance of colored patterns—pattern–color separability," J. Opt. Soc. Am. A **10**, 2458–2470 (1993).
26. L. E. Arend and A. Reeves, "Simultaneous color constancy," J. Opt. Soc. Am. A **3**, 1743–1751 (1986).
27. L. E. Arend, A. Reeves, J. Schirillo, and R. Goldstein, "Simultaneous color constancy: papers with diverse Munsell values," J. Opt. Soc. Am. A **8**, 661–672 (1991).
28. J. M. Troost and C. M. de Weert, "Naming versus matching in color constancy," Percept. Psychophys. **50**, 591–602 (1991).
29. E. W. Jin and S. K. Shevell, "Color memory and color constancy," J. Opt. Soc. Am. A **13**, 1981–1991 (1996).
30. M. P. Lucassen and J. Walraven, "Quantifying color constancy: evidence for nonlinear processing of cone-specific contrast," Vision Res. **33**, 739–757 (1993).
31. B. A. Wandell, *Foundations of Vision* (Sinauer, Sunderland, Mass., 1995).
32. P. K. Kaiser and R. M. Boynton, *Human Color Vision*, 2nd ed. (Optical Society of America, Washington, D.C., 1996).
33. J. S. Werner and J. Walraven, "Effect of chromatic adaptation on the achromatic locus: the role of contrast, luminance and background color," Vision Res. **22**, 929–944 (1982).
34. H. Helson, "Fundamental problems in color vision. I. The principle governing changes in hue, saturation and lightness of non-selective samples in chromatic illumination," J. Exp. Psychol. **23**, 439–476 (1938).
35. H. Helson and V. B. Jeffers, "Fundamental problems in color vision. II. Hue, lightness, and saturation of selective samples in chromatic illumination," J. Exp. Psychol. **26**, 1–27 (1940).
36. J. J. McCann, "Psychophysical experiments in search of adaptation and the gray world," *Proceedings of the IS&T's 47th Annual Conference* (Society for Imaging Science and Technology, Springfield, Va., 1994), pp. 397–401.
37. A. Valberg and B. Lange-Malecki, "'Color constancy' in Mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast," Vision Res. **30**, 371–380 (1990).
38. E. J. Breneman, "Corresponding chromaticities for different states of adaptation to complex visual fields," J. Opt. Soc. Am. A **4**, 1115–1129 (1987).
39. A. L. Gilchrist, "Perceived lightness depends on perceived spatial arrangement," Science **195**, 185 (1977).
40. A. L. Gilchrist, "When does perceived lightness depend on perceived spatial arrangements?" Percept. Psychophys. **28**, 527–538 (1980).
41. A. L. Gilchrist, S. Delman, and A. Jacobsen, "The classification and integration of edges as critical to the perception of reflectance and illumination," Percept. Psychophys. **33**, 425–436 (1983).
42. A. Gilchrist and A. Jacobsen, "Perception of lightness and illumination in a world of one reflectance," Perception **13**, 5–19 (1984).
43. E. H. Adelson, "Perceptual organization and the judgement of brightness," Science **262**, 2042–2044 (1993).
44. E. H. Adelson and A. P. Pentland, "The perception of shading and reflectance," in *Visual Perception: Computation and Psychophysics*, D. Knill and W. Richards eds. (Cambridge U. Press, New York, 1996).
45. W. S. Stiles, "Mechanism concepts in colour theory," J. Colour Group **11**, 106–123 (1967).

46. G. Buchsbaum, "A spatial processor model for object colour perception," J. Franklin Inst. **310**, 1–26 (1980).
47. L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectances," J. Opt. Soc. Am. A **3**, 29–33 (1986).
48. D. H. Brainard and W. T. Freeman, "Bayesian color constancy," J. Opt. Soc. Am. A **14**, 1393–1411 (1997).
49. CIE, *Colorimetry*, 2nd ed. (Bureau Central de la CIE, Paris, 1986).
50. I. Kuriki and K. Uchikawa, "Limitations of surface-color and apparent-color constancy," J. Opt. Soc. Am. A **13**, 1622–1636 (1996).
51. K. H. Bauml, "Color constancy in a Mondrian world," presented at the conference "From Genes to Perception," Tübingen, Germany, September 5–7, 1996.
52. K. H. Bauml, Institut fur Psychologie, Universität Regensburg, Regensburg, Germany, 1996 (personal communication).
53. D. H. Brainard, "Colorimetry," in *Handbook of Optics: Volume 1. Fundamentals, Techniques, and Design*, M. Bass ed. (McGraw-Hill, New York, 1995).
54. D. Katz, *The World of Colour* (Kegan, Paul, Trench Truber, London, 1935).
55. R. M. Evans, "Variables of perceived color," J. Opt. Soc. Am. **54**, 1467–1474 (1964).
56. R. M. Evans, *The Perception of Color* (Wiley, New York, 1974).
57. R. Mausfeld and R. Niederee, "An inquiry into relational concepts of colour, based on incremental principles of colour coding for minimal relational stimuli," Perception **22**, 427–462 (1993).
58. D. Jameson and L. Hurvich, "Essay concerning color constancy," Annu. Rev. Psychol. **40**, 1–22 (1989).
59. MATLAB *User's Guide* (The MathWorks, Inc., Natick, Mass., 1993).
60. A. Grace, *Optimization Toolbox for Use with* MATLAB—*User's Guide* (The MathWorks, Inc., Natick, Mass., 1990).
61. H. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," J. Opt. Soc. Am. A **3**, 1694–1699 (1986).
62. M. D'Zmura and G. Iverson, "Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces," J. Opt. Soc. Am. A **10**, 2148–2165 (1993).
63. M. D'Zmura and G. Iverson, "Color constancy. II. Results for two-stage linear recovery of spectral descriptions for lights and surfaces," J. Opt. Soc. Am. A **10**, 2166–2180 (1993).
64. D. A. Forsyth, "A novel algorithm for color constancy," Int. J. Comput. Vis. **5**, 5–36 (1990).
65. B. V. Funt and M. S. Drew, "Color constancy computation in near-Mondrian scenes using a finite dimensional linear model," IEEE Trans. Comput. Vis. Pattern Recog. (1988).
66. B. V. Funt and M. S. Drew, "Color space analysis of mutual illumination," IEEE Trans. Pattern. Anal. Mach. Intell. **15**, 1319–1326 (1993).
67. G. D. Finlayson, "Color constancy in diagonal chromaticity space," in *Proceedings of the 5th International Conference on Computer Vision* (IEEE Computer Society Press, Los Alamitos, Calif., 1995), pp. 218–223.
68. G. D. Finlayson and B. V. Funt, "Color constancy under varying illumination," in *Proceedings of the 5th International Conference on Computer Vision* (IEEE Computer Society Press, Los Alamitos, Calif., 1995), pp. 720–725.
69. H. J. Trussell and M. J. Vrhel, "Estimation of illumination for color correction," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* (IEEE, New York, 1991); pp. 2513–2516.
70. M. D'Zmura, G. Iverson, and B. Singer, "Probabilistic color constancy," in *Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow Indow's 70th Birthday*, R. D. Luce, M. D'Zmura, D. Hoffman, G. Iverson, and A. K. Romney eds. (Erlbaum, Mahwah, N. J., 1995).
71. W. T. Freeman and D. H. Brainard, "Bayesian decision theory, the maximum local mass estimate, and color constancy," in *Proceedings of the 5th International Conference on Computer Vision* (IEEE Computer Society Press, Los Alamitos, Calif., 1995), pp. 210–217.
72. J. M. Speigle and D. H. Brainard, "Luminosity thresholds: effects of test chromaticity and ambient illumination," J. Opt. Soc. Am. A **13**, 436–451 (1996).
73. D. H. Brainard, M. D. Rutherford, and J. M. Kraft, "Color constancy compared: experiments with real image and color monitors," Invest. Ophthalmol. Visual Sci. Suppl. **38**, S476 (1997).
74. For the perceived-surface-color task, which we do not believe our observers were performing, the mean constancy index from Arend *et al.*[27] was 0.44, still lower than ours. In recent experiments, however, Bauml[51] has found a mean constancy index of 0.78 for a perceived-surface-color task.
75. V. Smith and J. Pokorny, " Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," Vision Res. **15**, 161–171 (1975).
76. P. DeMarco, J. Pokorny, and V. C. Smith, "Full-spectrum cone sensitivity functions for X-chromosome-linked anomalous trichromats," J. Opt. Soc. Am. A **9**, 1465–1476 (1992).
77. K. L. Kelly, K. S. Gibson, and D. Nickerson, "Tristimulus specification of the *Munsell Book of Color* from spectrophotometric measurements," J. Opt. Soc. Am. **33**, 355–376 (1943).
78. D. Nickerson, "Spectrophotometric data for a collection of Munsell samples," U.S. Dept. of Agriculture Rep. (U.S. Government Printing Office, Washington, D.C., 1957; available from Munsell Color Company, Baltimore, Md.).
79. D. Nickerson and D. H. Wilson, "Munsell reference colors now specified for nine illuminants," *Illum. Eng.* **45**, 507–517 (1950).
80. J. M. Speigle and D. H. Brainard, "Is color constancy task independent?" in *Proceedings of the 4th Information Science and Technology/Society for Information Display Color Imaging Conference: Color Science, Systems, and Applications* (Society for Imaging Science and Technology, Springfield, Va., 1996), pp. 167–172.
81. D. H. Brainard and W. A. Brunt, "The equivalent illuminant," Invest. Ophthalmol. Visual Sci. Suppl. **37**, S648 (1996).
82. J. M. Speigle, W. A. Brunt, and D. H. Brainard, "Color constancy measured using three tasks," Invest. Ophthalmol. Visual Sci. Suppl. **37**, S1064 (1996).