The color constancy of three-dimensional objects

Bei Xiao
Graduate Program in Neuroscience, University of Pennsylvania, Philadelphia, PA, USA, & Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Brendan Hurst
Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA, & Department of Psychological and Brain Sciences, The Johns Hopkins University, Baltimore, MD, USA

Lauren MacIntyre
Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA, & Brain Imaging and EEG Laboratory, University of California San Francisco, San Francisco, CA, USA

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

Human color constancy has been studied for over 100 years, and there is extensive experimental data for the case where a spatially diffuse light source illuminates a set of flat matte surfaces. In natural viewing, however, three-dimensional objects are viewed in three-dimensional scenes. Little is known about color constancy for three-dimensional objects. We used a forced-choice task to measure the achromatic chromaticity of matte disks, matte spheres, and glossy spheres. In all cases, the test stimuli were viewed in the context of stereoscopically viewed graphics simulations of three-dimensional scenes, and we varied the scene illuminant. We studied conditions both where all cues were consistent with the simulated illuminant change (consistent-cue conditions) and where local contrast was silenced as a cue (reduced-cue conditions). We computed constancy indices from the achromatic chromaticities. To first order, constancy was similar for the three test object types. There was, however, a reliable interaction between test object type and cue condition. In the consistent-cue conditions, constancy tended to be best for the matte disks, while in the reduced-cue conditions constancy was best for the spheres. The presence of this interaction presents an important challenge for theorists who seek to generalize models that account for constancy for flat tests to the more general case of three-dimensional objects.

Keywords: color constancy, color vision, shape and color

Citation: Xiao, B., Hurst, B., MacIntyre, L., & Brainard, D. H. (2012). The color constancy of three-dimensional objects. *Journal of Vision, 12*(4):6, 1–15, http://www.journalofvision.org/content/12/4/6, doi:10.1167/12.4.6.

Introduction

Color constancy is classically conceived as the ability to perceive stable object colors, despite variation of illumination and scene context. Such constancy is difficult to achieve because the light reflected from objects varies with the illumination. Given that there is considerable variation in the spectral and spatial properties of natural illuminants, constancy helps object color to be a useful guide to object identity.

Many studies of color constancy have examined the color appearance of diffusely illuminated flat matte surfaces (e.g., Arend & Reeves, 1986; Brainard, 1998; Brainard, Brunt, & Speigle, 1997; Brainard & Wandell, 1992; Delahunt & Brainard, 2004; Granzier, Brenner, Cornelissen, & Smeets, 2005; Helson & Jeffers, 1940; Helson & Michels, 1948; Kraft & Brainard, 1999; McCann, McKee, & Taylor, 1976; Olkkonen, Hansen, & Gegenfurtner, 2009). Taken as a whole, these studies reveal that under many conditions the human visual system exhibits considerable constancy, and some of the computations and mechanisms that support such constancy have been delineated (reviews in Brainard, 2004; Brainard & Maloney, 2011; Foster, 2011).

Constancy has also been studied for flat matte surfaces viewed in spatially rich three-dimensional scenes. Embedding stimuli in three-dimensional scenes allows one to study the effect of three-dimensional surface pose on perceived surface lightness and color (Bloj, Kersten, &

doi: 10.1167/12.4.6 Received August 20, 2011; published April 16, 2012 ISSN 1534-7362

Hurlbert, 1999; Boyaci, Doerschner, & Maloney, 2004; Boyaci, Maloney, & Hersh, 2003; Ripamonti et al., 2004; Werner, 2006; see also Gilchrist, 1977, 1980; Hochberg & Beck, 1954), as well as the effects of cues to the illuminant (e.g., specular highlights on surrounding objects, mutual illumination between surfaces) that are not available when all surfaces in the scene are flat, co-planar, and matte (Doerschner, Boyaci, & Maloney, 2004; Kraft, Maloney, & Brainard, 2002; Yang & Maloney, 2001; Yang & Shevell, 2003; see also Funt & Drew, 1993; Lee, 1986; Tominaga & Wandell, 1989). Some theoretical approaches to understanding the effects of three-dimensional geometry for flat test surfaces are available (Brainard & Maloney, 2011; Gilchrist, 2006; Gilchrist et al., 1999).

In natural scenes, however, only a small fraction of objects are flat and matte. Indeed, the reflectance of most real objects must be characterized as a function of the direction of incident and reflected light rays. Such a characterization is called the object's bidirectional reflectance distribution function (BRDF). Object BRDFs vary according to what the object is made from: different materials such as plastic, metal, wax, and fabric have different BRDFs. How the material properties of three-dimensional objects interact with their perceived lightness and color, as well as how we perceive the material properties of objects, is receiving increased attention (Blake & Bülthoff, 1990; Fleming & Bülthoff, 2005; Fleming, Dror, & Adelson, 2003; Kim, Marlow, & Anderson, 2011; Marlow, Kim, & Anderson, 2011; Motoyoshi, Nishida, Sharan, & Adelson, 2007; Nishida & Shinya, 1998; Olkkonen & Brainard, 2010, 2011; Pessoa, Mingolla, & Arend, 1996; Sharon, Li, Motoyoshi, Nishida, & Adelson, 2008; Todd, Norman, & Mingolla, 2004; Xiao & Brainard, 2008).

Little is known about the color constancy of three-dimensional objects, particularly those made of non-matte materials (but see de Almeida, Fiadeiro, & Nascimento, 2010; Hedrich, Bloj, & Ruppertsberg, 2009; Olkkonen & Brainard, 2010, 2011). Here we report experiments that address this issue. We used a forced-choice procedure to determine the chromaticities at which matte and glossy spheres, embedded in three-dimensional scenes, appeared achromatic. We also determined such achromatic chromaticities for flat matte disks embedded in the same scenes. We varied the scene illuminant and the reflectance of the surface immediately behind the tests and used the results to determine constancy indices.

A focus of the experiments is the question of whether the constancy found for three-dimensional spheres bears a simple relation to the constancy found for flat matte tests. More specifically, we asked whether there is an interaction between the test object's three-dimensional shape and spectral changes in the contextual scene. Answering this question is key to determining whether extant theories of constancy, developed for flat matte tests, may be easily generalized to handle three-dimensional objects.

Experiment 1

Methods

Apparatus

Stimuli were presented to observers stereoscopically using two CRT monitors (Hewlett-Packard, Palo Alto, CA, Model p1230). The monitors were driven at a 75-Hz refresh rate at 1152 by 870 spatial resolution and with a color lookup table providing 14-bit (pseudocolor) resolution for each color channel (BITS++, Cambridge Research Systems, Rochester, England). Two small mirrors were placed directly before the observers. The mirrors reflected the images displayed on the left and right monitors to the left and right eyes. Observers viewed the monitors through two square apertures, one for each eye. Black poster board and black cloth were used liberally to minimize stray light reaching the observer.

We used custom software to align the images from the two monitors. The mirrors in the display were temporarily replaced by beam splitters, allowing a common reference grid to be visually superimposed with each monitor. An observer manually positioned a dot displayed on each monitor to align with each grid intersection. We used the positions of the aligned dots to determine transformations between image and monitor coordinates, so that during the experiment each displayed image was mapped to the grid-aligned coordinate system. The distance from the cyclopean eye to the alignment grid was 76.4 cm. This was also the approximate optical distance of each monitor to the corresponding eye.

Stimulus scenes

The stimuli consisted of graphics simulations of illuminated scenes. The geometry of the scenes was created using the Maya software tools (Alias, San Rafael, CA). They were rendered using the Radiance package (Ward, 1994; http://radsite.lbl.gov/radiance/HOME.html), augmented by custom Matlab tools (Render Toolbox; http://rendertoolbox.org) that allowed us to specify spectral surface reflectance functions and illuminant spectral power distributions (400–700 nm at 10-nm intervals) as well as to render on a wavelength-by-wavelength basis. This process resulted in a 31-plane hyperspectral image of each scene. The hyperspectral images were converted to three-plane LMS representations by computing at each pixel the excitations that would be produced in the human L-, M-, and S-cones. We used the Stockman–Sharpe cone fundamentals for this calculation (CIE, 2007; Stockman & Sharpe, 2000).

Scenes were rendered from two viewpoints, corresponding to the positions of the observer's two eyes (6-cm horizontal separation). Each eye received the image rendered from the corresponding viewpoint.

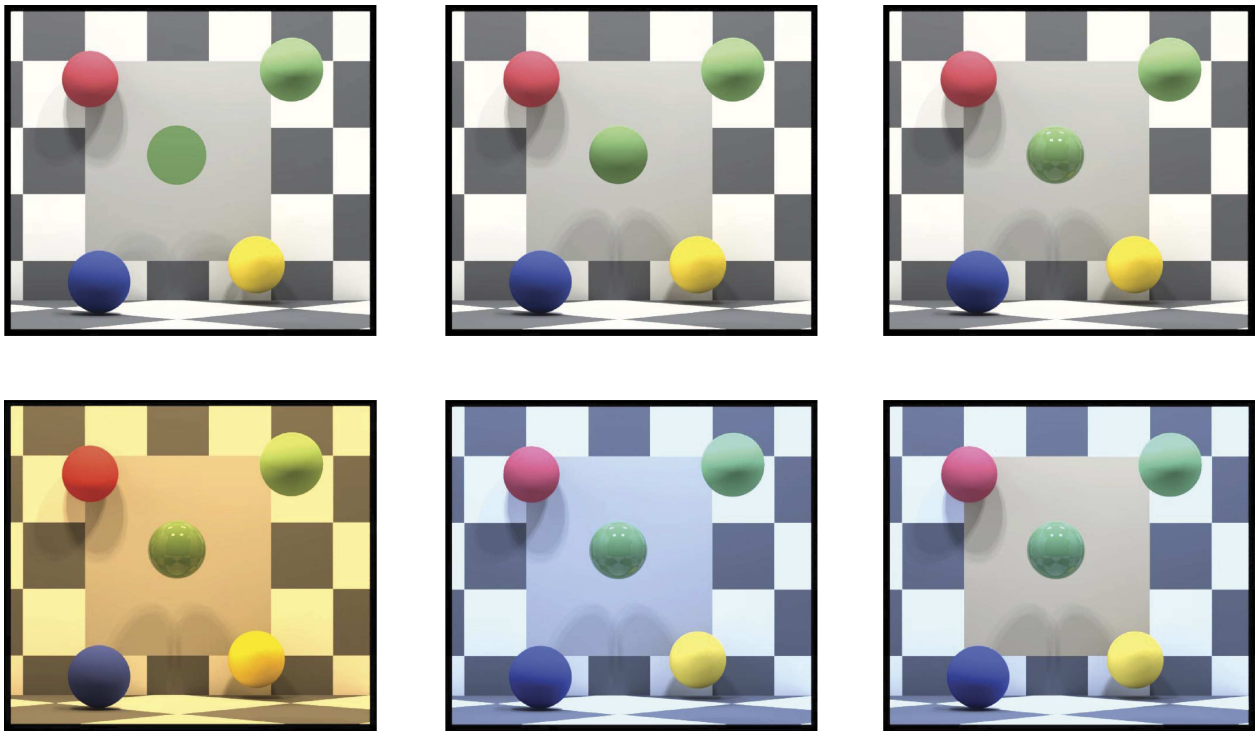

Figure 1. Examples of experimental images. The three images in the top row show the experimental scene rendered using the Neutral illuminant. Across these three images, the test object varies. (Left) Matte disk. (Center) Matte sphere. (Right) Glossy sphere. The three images in the bottom row show the rendered scene for the Yellow illuminant consistent-cue condition (left), Blue illuminant consistent-cue condition (center), and Blue illuminant reduced-cue condition (right). The images shown in the figure were converted from the LMS cone representation to the sRGB color space for display. They were also tone mapped using an automated procedure that scaled each image so that its maximum luminance value was five times the image mean and then clipped out-of-range pixels. This tone-mapping procedure differs from that applied to the experimental images and was used here for convenience.

Each scene contained a test object whose color appearance was judged by observers. The test object could be a flat matte disk, a matte sphere, or a glossy sphere. These three test objects are illustrated by the images shown in the top row of Figure 1.

The rest of the scene provided the context in which the test object was viewed. The immediate surround of the test object consisted of a spatially uniform background on the back wall of the rendered scene. The global surround was provided by a checkered pattern on the remainder of the back wall and floor of the scene, as well as by four contextual spheres (see Figure 1). A summary of key scene parameters is provided below.

We specify *xyz* scene coordinates in centimeters relative to an origin defined by the position of the cyclopean eye of the observer. The x coordinate decreases to the observer's left and increases to the right; the y coordinate decreases in the down direction and increases in the up direction; the z coordinate increases with increasing distance from the observer. In this coordinate system, the center of the test objects was located at (0.2, 2.3, 72.1) cm and the diameters of the test objects were 3.3 cm (disks) and 3.2 cm (spheres).

The center of the spatially uniform region background behind the test object was located at (0.4, 2.1, 76.3) cm and had xy dimensions of 10.8 by 11.7 cm. The floor was located at a y coordinate of −6 cm. The four contextual spheres all had a diameter of 3.2 cm and had centers located at (clockwise from upper left) (−4.7, 6.7, 74.7), (5.7, 6.5, 66.4), (4.8, −3.8, 74.2), and (−4.1, −4.4, 67.3) cm. The front wall of the scene was at a z coordinate of 55.4 cm. It contained an aperture through which the scene was visible. The xy center position of the aperture was (0.0, 0.0) cm and its xy size was 15.2 by 13.5 cm. Not visible through the aperture were side walls at x positions of −16.6 and 16.9 cm and a ceiling at a y position of 27.8 cm. The interior of the front wall, the side walls, and the ceiling all had the same spatially uniform reflectance.

The geometric surface reflectance properties of all objects in the scenes were specified with respect to the isotropic Ward reflection model (Ward, 1992), as implemented in Radiance using its plastic material type. The RenderToolbox allowed us to vary the diffuse reflectance component of objects as a function of wavelength; the non-diffuse (specular) reflectance component was constant as a function of wavelength. In Experiment 1, all surfaces were matte (diffuse reflectance only), except for the glossy test sphere. The Ward model parameters for non-diffuse reflectance component of this sphere were 0.02 (gloss parameter) and 0.00 (roughness parameter) at

all wavelengths. No correction was made for the fact that before rendering, Radiance applies a normalization to the magnitude of the specified diffuse reflectance component of glossy objects, so that the diffuse component of the matte and glossy test objects varied slightly in overall magnitude. The diffuse spectral reflectance functions of the contextual surfaces in the scene are provided in the Supplementary material.[1] These were held fixed across conditions, with the exception of the uniform surface that provided the immediate background for the test object. This was varied according to condition as described in more detail below.

Each scene was illuminated by four area light sources and one diffuse light source. The area lights were parallel to the floor of the scene and each had *xz* extents of 4 cm by 4 cm. The center coordinates of the four area lights were (7.0, 26.0, 77.4), (−7.0, 26.0, 77.4), (7.0, 26.0, 57.9), and (−7.0, 26.0, 57.9) cm. For each illuminant condition, all light sources shared the same spectrum. Three illuminant spectra were used. The Neutral illuminant was CIE daylight D65 (CIE, 2004). The Blue and Yellow illuminants were constructed from the CIE linear model for daylights (CIE, 2004; Judd, MacAdam, & Wyszecki, 1964). The Blue illuminant had a correlated color temperature of 15,000 K and the Yellow illuminant had a correlated color temperature of 3500 K. The spectra of these illuminants are provided in the Supplementary material. Scenes rendered under the Yellow and Blue illuminants are shown in the bottom row of Figure 1 (left and center images).

The illuminant intensities were specified in arbitrary units. The resulting images were then scaled by a common (across conditions) factor to bring them into the general intensity range of our monitors. To characterize the overall intensity of each scene, we measured the mean pixel luminance values of the uniform background surface behind the test object. These values are provided in the Supplementary material, as part of the data tabulated for each individual observer.

To examine the role of local contrast on constancy, we studied both consistent-cue and reduced-cue conditions. In the consistent-cue conditions, only the illuminant was varied, while the contextual surfaces in the scene remained fixed. In the reduced-cue conditions, we co-varied the simulated reflectance of the central patch of the background wall with the illuminant, so as to hold the light reflected from that part of the wall constant. This manipulation tends to reduce constancy (Brainard, 1998; Delahunt & Brainard, 2004; Kraft & Brainard, 1999) and can provide data diagnostic for testing models of constancy (Brainard et al., 2006; Brainard & Maloney, 2011). We refer to the immediate background of the test object in the consistent-cue conditions as the normal background; we refer to the immediate background of the test object in the reduced-cue condition as the equated background. The reflectance functions used for the background central patch for each condition are provided in the Supplementary material. The center and right images in the bottom row of Figure 1 illustrate the difference between consistent-cue and reduced-cue conditions.

Because some of the rendered images had a high-dynamic range, it was necessary to tone map the images after the overall scaling described above had been applied. We used a simple tone-mapping procedure that truncated pixel luminance to 25 cd/m^2 while holding the chromaticity of each truncated pixel constant.

The scaled and tone-mapped LMS images were converted for display using standard methods that accounted for the spectral power distributions of the monitor's phosphors and for the monitor's gamma functions (Brainard, Pelli, & Robson, 2002).

Due to a bug in the experimental software, the images were displayed on the experimental monitors at 80% of their intended size. (This bug was corrected for Experiment 2.)

Manipulation of test object color

To vary the cone coordinates of the test objects without full re-rendering of the scene, we pre-rendered each test object with eight basis surface reflectances. We could then recombine the pre-rendered images at run time to produce test objects with any desired colorimetric coordinates (see Xiao & Brainard, 2008; also Griffin, 1999). The reported *xy* chromaticities and luminances of test objects were obtained by taking the spatial average of the Judd–Vos (Vos, 1978) *XYZ* tristimulus coordinates of the pixels in the image of the object.

Summary of conditions

There were two shapes for the test object (sphere and disk). For the spheres, there were two materials, matte and glossy. For each test object, there were five contextual conditions (Neutral, consistent-cue Blue, consistent-cue Yellow, reduced-cue Blue, and reduced-cue Yellow).

Task

A single condition was studied in each experimental session, with multiple sessions typically run on each day with short breaks between sessions. Observers ran in at least one practice session to become familiar with the experimental procedures before starting on experimental sessions.

We used a forced-choice procedure to determine the LMS coordinates of test objects that appeared achromatic (Chichilnisky & Wandell, 1999). At the start of an experimental session, observers were presented with a scene of an empty room and initiated a block of trials by pressing a button. On each trial, a test object appeared for 0.75 s in the scene and then disappeared, and observers performed one of two forced-choice tasks. On *red–green*

trials, observers pressed one button if the test object appeared reddish, and another if it appeared greenish. On *blue–yellow* trials, observers pressed one button if the test appeared bluish and another if it appeared yellowish. The two types of trials were interleaved in blocks of seven trials each. At the beginning of each block, a speech synthesizer announced the trial type for the next seven trials. We discarded the first trial after each staircase-type switch to minimize contamination caused by observer confusion about which judgment was being performed. An experimental session consisted of 15 blocks of each trial type (180 non-discarded trials per session).

Across blocks, the trials were organized into red–green and blue–yellow staircase pairs that sought the achromatic point for the test object. Three staircase pairs were run in each session, each with a different randomly chosen starting point (see below). Each member of a staircase pair consisted of 30 trials, for a total of 180 non-discarded trials per session. Within each block, two trials were run for the corresponding member of each of the three staircase pairs. The order of trials for the three staircase pairs was randomized in subblocks of three trials each.

The procedure for each staircase pair is most easily understood by reference to Figure 2A, which shows data from one staircase pair in an (r, b) chromaticity space defined with respect to the monitor primaries. Let R, G, and B represent the linear weights of each monitor primary required to produce the average LMS coordinates of the test object. Then

$$r = R/(R + G + B), \quad (1)$$

$$b = B/(R + G + B). \quad (2)$$

Specification of the (r, b) coordinates of the test object, together with its luminance, is sufficient to reconstruct the average LMS coordinates of the test object and vice versa. The average luminance of the test objects was held constant at 5 cd/m^2.

To initialize each staircase pair, (r, b) coordinates were chosen according to a uniform pseudorandom distribution within the range [0, 1]. If the result corresponded to *RGB* coordinates for the test object that were within the monitor gamut, the chosen (r, b) coordinates were used as the starting point for the staircase pair. If not, new (r, b) coordinates were drawn and the process repeated until a within-gamut starting point was found. Once initialized, the observer's responses were used to modify the r and b values of the test object. On red–green trials, a reddish response led to a decrease of the r chromaticity of the test object while the b value was left unchanged; a greenish response led to an increase of the r value, again with the b value unchanged. On blue–yellow trials, a blue response led to a decrease of the b value with the r value unchanged and a yellowish response led to an increase of the b value with the r value unchanged. The initial step size of the adjustments was 0.06 for both r and b. After the second reversal, the step size decreased to 0.03. After the fourth reversal, it decreased again to 0.014. A final decrease to 0.006 occurred after the fifth reversal. Thereafter, the adjustment step size was held constant. For each staircase pair, the r and b coordinates of the achromatic point were obtained from the corresponding staircase of the pair. These were taken as the mean of all reversals when both staircases had reached the smallest step size. We took the value for a reversal to be the average of the value on the reversal trial and on the preceding trial.

Figure 2A shows the data for one staircase pair from one session. Note that the step size decreases at several of the reversal points and that the staircases converge well before the 30th trial. Figure 2B plots the trajectories of all three staircase pairs from this session. Note that although the staircases begin at different points in the (r, b) chromaticity space, they converge to a common point.

Although the staircases generally converged well, we implemented procedures to exclude bad staircase pairs and sessions. First, to verify that a staircase pair had converged we required that both staircases in a pair had reversed at least once after each had reached its smallest step size. Staircase pairs that did not meet this convergence criterion were dropped. To include data from a session, we required that at least two of the staircase pairs within that session converged by the above criterion. For such sessions, we took the r and b values of the achromatic point to be the mean r and b values from the staircase pairs that had converged. If the within session standard deviation for either the r or b values from the staircase pairs that had converged exceeded 0.02, we excluded the data from that session from further analysis.

Observers

Four naive observers (TYD, female, age 18 years; YRT, male, 19 years; SXT, female, 22 years; FJK, female, 20 years) completed the experiment. Each had corrected Snellen acuity of 20/30 or better in both eyes, normal color vision as assessed by the Ishihara plates (Ishihara, 1977; no more than 2 plates incorrect), and normal stereo vision as assessed by the Keystone VS-II Model 1135A vision screening system (no more than 1 incorrect response). For observer FJK, however, we discovered at the end of the experiment that the latter 2/3 of her sessions failed to meet our inclusion criteria. This observer's data are not included in the analyses presented in the paper but, for completeness, are available in the Supplementary material. Two additional observers started the experiment. One was discontinued after running in three experimental sessions because his staircases did not reliably converge. The other was discontinued after running in nine experimental

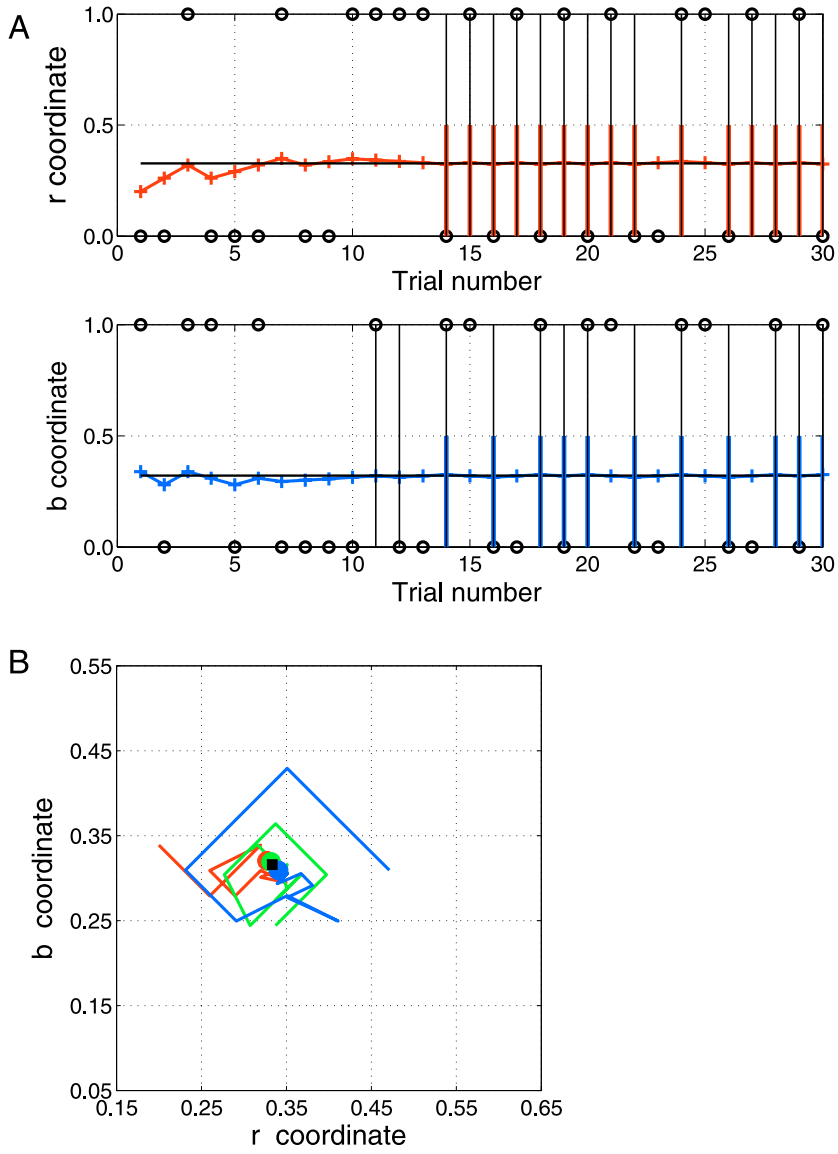

Figure 2. (A) Data from one staircase pair (observer SXT, matte disk, Neutral illuminant). The top plot shows the test object *r* value on every red–green trial while the bottom plot shows the test object *b* value on every blue–yellow trial. Open circles indicate the observer's response on each trial (a circle plotted at 0 indicates green and yellow responses, while a circled plotted at 1 indicates red and blue responses). Vertical black lines extending between 0 and 1 indicate reversals at the smallest step size. Vertical colored lines extending between 0 and 0.5 indicate reversals averaged in computation of achromatic point for this staircase pair. Both staircases had to have reached the smallest step size before reversals for either staircase were averaged. In applying this criterion, we neglected effects of trial-type blocking and took as concurrent trials from each staircase type that had the same trial number within their respective staircases. (B) Trajectory of the three staircase pairs from one session. Same session as for staircase pair shown in (A). Each colored line indicates the trajectory of (*r*, *b*) chromaticities for one staircase pair. The *b* value for each trial of the *b* staircase is plotted against the *r* value for the corresponding trial of the *r* staircase. Because the two trial types were interleaved in blocks, these are not the exact (*r*, *b*) chromaticities shown on any particular trial; the trajectories shown, however, provide a visual sense of the progression of each staircase pair. The closed circles indicate the (*r*, *b*) chromaticity of the achromatic locus determined from each staircase pair, and the black square shows the average taken over the three staircases.

sessions for failing to reliably appear for scheduled experimental sessions. Data from these two observers was not analyzed further.

Each observer completed three sessions for each experimental condition (test object, illuminant, and cue condition), with one exception. Due to experimenter error, observer SXT completed four sessions for the matte disk/blue illuminant/equated background condition and two sessions for the matte sphere/blue illuminant/equated background condition. After application of our exclusion criteria, data from at least two sessions were available for all conditions for all observers.

The experimental procedures conformed to the tenets of the Declaration of Helsinki.

Instructions

Observers were instructed as to the experimental procedures and the perceptual judgments they would be asked to make. The instructions emphasized that we wanted observers to judge the surface properties of the test objects, along the lines of the "paper instructions" used by Arend and Reeves (1986). The full instructions provided to observers are available as part of the Supplementary material.

In order to familiarize observers with the type of appearance judgment we wanted them to make, observers went through an induction procedure (Allred & Brainard, 2009; Doerschner, Boyaci, & Maloney, 2007) prior to receiving the experimental instructions. Here observers viewed a grayscale test cube placed in an illuminated chamber. Observers were asked to match the paint of the test surface on the cube to that of a patch on a grayscale palette seen in an adjoining chamber. Observers were asked to make such matches for two additional cubes, with two other grayscale surface colors. After the matches were done, observers were allowed to handle three test blocks for up to 2 min. They then performed an additional short series of grayscale paint matches. The induction procedure was used to familiarize the observer with judging the appearance of object surfaces, and the responses were not analyzed. Additional details about the induction procedure are available as part of the Supplementary material.

Results

For all three test objects, constancy was good for the consistent-cue condition. This is shown for Observer SXT by the filled circles in each panel of Figure 3, which show the achromatic chromaticities for the consistent-cue conditions (the three panels show data for the disk, matte sphere, and glossy sphere, respectively.) Between the Neutral condition and the consistent-cue Blue and Yellow conditions, the achromatic chromaticities approximately track the changes in illuminant chromaticities (shown in the figure by the crosses). This is the pattern we expect in the data for a color constant visual system (Brainard, 1998; Delahunt & Brainard, 2004).

In the reduced-cue conditions, constancy was much reduced. This is shown for Observer SXT by the open circles in each panel of Figure 3, which show the achromatic chromaticities for the reduced-cue Blue and Yellow conditions. These chromaticities lie much nearer to that of the Neutral illuminant than for the corresponding consistent-cue data and indicate a smaller degree of constancy. This observation is in accord with previous studies (Brainard, 1998; Delahunt & Brainard, 2004; Kraft & Brainard, 1999).

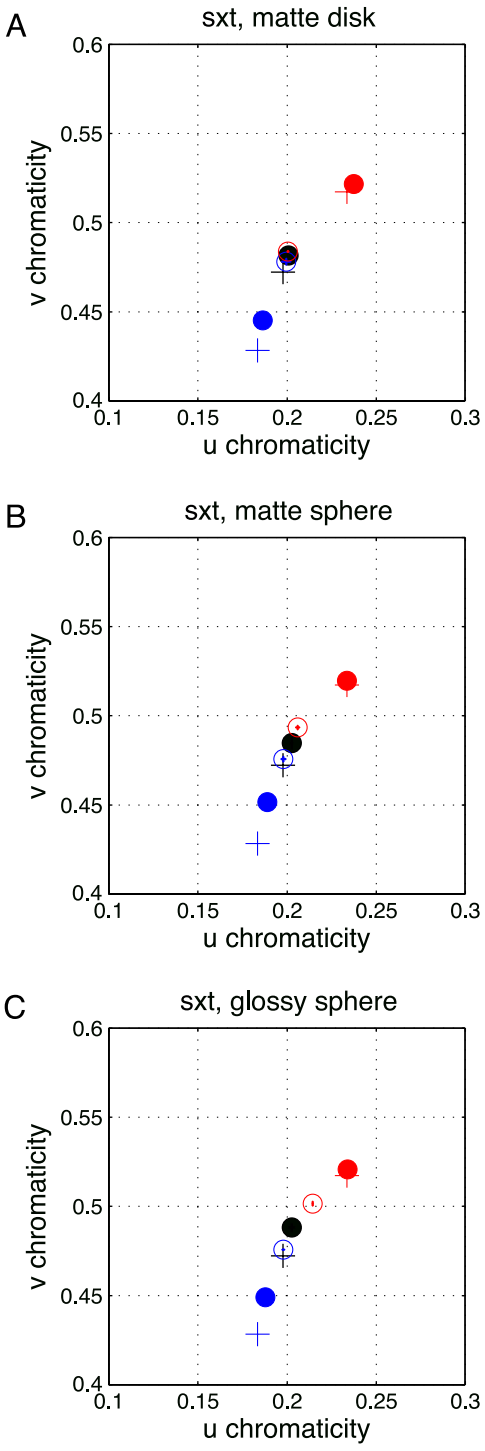

Figure 3. Achromatic chromaticities, Experiment 1, Observer SXT. (A) Matte disk. Plus symbols represent illuminant chromaticities. Solid circles represent consistent-cue condition achromatic chromaticities. Open circles represent reduced-cue achromatic chromaticities. Error bars are ±1 *SEM*. Black symbols are for Neutral illuminant, blue symbols are for Blue illuminant, and red symbols are for the Yellow illuminant. (B, C) Matte sphere and glossy sphere, respectively. Same format as (A). Plots in the same format for the other observers, as well as tabulated individual observer data, are provided in the Supplementary material.

We used the achromatic chromaticities to compute a constancy index for each illuminant change and condition (see Delahunt & Brainard, 2004). The index takes on a value of 1 for perfect constancy and a value of 0 for no constancy and was computed as follows. First, we used the achromatic chromaticity for the Neutral condition together with the spectral power distribution of the Neutral illuminant to compute an equivalent achromatic surface reflectance function. This is a surface reflectance function whose chromaticity under the Neutral illuminant matches the achromatic chromaticity. We then used the equivalent achromatic surface reflectance function to compute an equivalent illuminant spectral power distribution for each additional condition, so that the chromaticity of the equivalent achromatic surface reflectance function under the equivalent illuminant matched the measured achromatic chromaticity for each condition. We then found the uv chromaticity of the equivalent illuminant and computed the constancy index as follows:

$$CI = 1 - \|uv_{eq} - uv_{illum}\| / \|uv_{neutral} - uv_{illum}\|, \quad (3)$$

where uv_{eq} is a vector specifying the CIE uv chromaticity of the equivalent illuminant, uv_{illum} is a vector specifying the chromaticity of the changed (i.e., either Blue or Yellow) illuminant, and $uv_{neutral}$ is a vector specifying the chromaticity of the Neutral illuminant. In computing the equivalent achromatic surface reflectance function and equivalent illuminants, we used three-dimensional linear models for surface and illuminant spectral power distributions. The surface linear model was computed from measured surface reflectance functions of the Munsell papers (Newhall, Nickerson, & Judd, 1943; Nickerson, 1957) and the illuminant linear model was the CIE daylight model (CIE, 2004). Because they are computed from the equivalent illuminants, our constancy indices are largely insensitive to any overall shifts in appearance with test object shape that occur within the Neutral illuminant condition and thus allow us to focus our attention on constancy with respect to our contextual manipulations.

Figure 4A shows the constancy indices for Observer SXT, computed from the achromatic chromaticities shown in Figure 3. Figure 4B shows the mean constancy indices for all three observers. Several features of the data emerge from the constancy indices. First, the indices capture the good constancy for the consistent-cue conditions and decreased constancy for the reduced-conditions that was apparent in the achromatic chromaticities shown in Figure 3.

Second, there is an interaction between test object type and constancy. For the consistent-cue conditions, constancy is best for the matte disk, while for the reduced-cue conditions constancy is worst for the matte disk. While modest in size, the effects of test object are statistically significant (see Table 1). Examination of the pattern of variation in the indices suggests that the change from matte disk to matte sphere interacts with cue condition

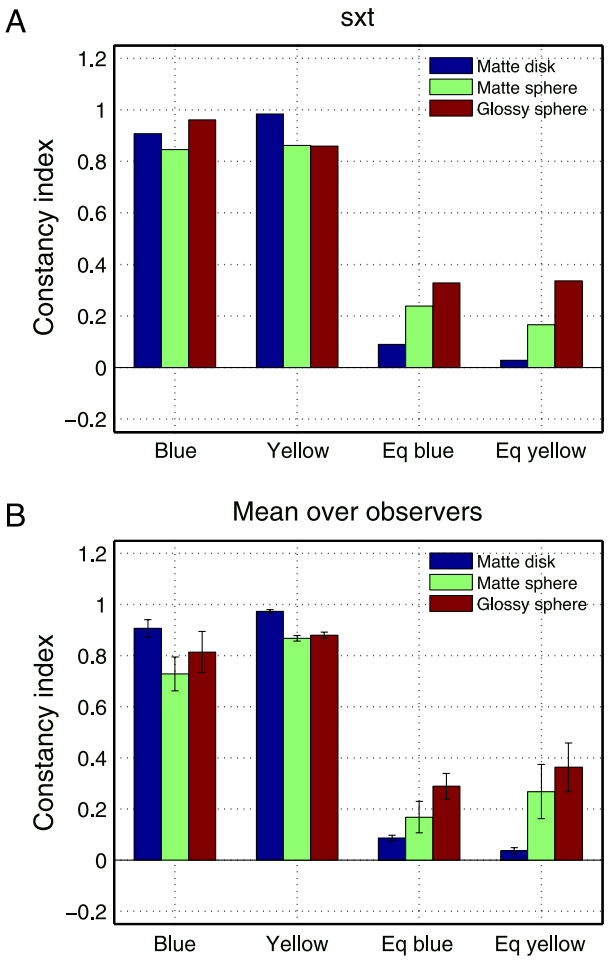

Figure 4. Constancy indices, Experiment 1. (A) Constancy indices for Observer SXT. Each group provides indices for one condition. Within group, the three bars provide indices for the matte disk, matte sphere, and glossy sphere. Blue: Blue illuminant change, consistent cue; Yellow: Yellow illuminant change; consistent cue; Eq Blue: Blue illuminant change, reduced cue; Eq Yellow: Yellow illuminant change, reduced cue. Constancy index plots for the other observers are available as part of the Supplementary material. (B) Mean constancy indices. Same format as (A). Error bars show ±1 *SEM*.

and that there is an improvement of constancy with change from matte sphere to glossy sphere. The improvement of constancy with the addition of specular highlights is consistent with previous reports (Yang & Maloney, 2001; Yang & Shevell, 2003).

Experiment 2

Purpose

The interaction between cue condition and test object found in Experiment 1 is important, since such an

Source	VarType	SumSq	*df*	MeanSq	*F*	*p*
Observer	Random	0.04	2	0.02	2.52	n.s. (0.10)
Test object	Fixed	0.05	2	0.03	3.26	n.s. (0.06)
Cue condition	Fixed	3.93	1	3.93	469	<0.001
Illuminant	Fixed	0.04	1	0.04	4.68	<0.05
Test object * Cue condition	Fixed	0.22	2	0.11	13.3	<0.001
Test object * Illuminant	Fixed	0.02	2	0.01	1.13	n.s. (0.34)
Cue condition * Illuminant	Fixed	0.01	1	0.01	0.65	n.s. (0.43)
Error	Random	0.2	24	0.01		
Total		4.51	35			

Table 1. ANOVA for Experiment 1. Observers were coded as a random variable. The main effect of each variable and the two-way interactions between all variables except observer were modeled.

interaction means that generalizing results obtained with flat tests to constancy for three-dimensional objects cannot be accomplished simply by measuring the effect of object shape on color appearance under a single illuminant. The main purpose of Experiment 2 was to determine whether this effect replicated with a new set of observers and to explore the effects of other stimulus manipulations. In particular, an important difference between the disk and matte sphere is that the disk was essentially homogeneous in its luminance, while there was luminance variation across the sphere. In addition, the test luminance used in Experiment 1 was very close to that of the immediate surround that reduced the visibility of the border dividing the background from the matte disk. We manipulated test luminance in Experiment 2 to determine whether the effect of shape could be mimicked by overall test luminance variation and to include a condition where the background–disk border was not close to isoluminant. We also explored the effect of adding specular highlights to the contextual spheres. Since in Experiment 1 the effect of illuminant change was small and since there were no interactions with the illuminant (Table 1), in Experiment 2 we studied only the Blue illuminant change.

Methods

The methods for Experiment 2 were the same as those for Experiment 1 with the following changes:

- Only the Blue illuminant change was studied.
- Test luminances of both 3 cd/m^2 and 5 cd/m^2 were used.
- A condition in which the contextual spheres were glossy rather than matte was added. The Ward parameters for these spheres matched those of the glossy test sphere: glossiness = 0.02, roughness = 0.00. This condition was run only for the 5 cd/m^2 tests.
- The software bug noted in the methods for Experiment 1 was fixed, so that the rendered images were displayed at their intended size.
- Five new naive observers participated (VJC, female, age 23 years; DYW, male, age 30 years; JOP, male, age 24 years; HBP, female, 24 years; UOS, female, age 18 years). Each had corrected Snellen acuity of 20/40 or better in both eyes, normal color vision as assessed by the Ishihara plates (Ishihara, 1977; no more than 2 plates incorrect), and functional stereo vision as assessed either by the Keystone VS-II Model 1135A vision screening system (no more than 1 incorrect response) or by a custom program run on our apparatus, in which they discriminated whether a square defined by binocular disparity appeared in front of or behind a spatially uniform background. One additional observer ran in six experimental sessions but was discontinued because her staircase data did not reliably converge. Data from this observer were not analyzed further.
- The number of trials per staircase was reduced to 20, and the number of target sessions per condition reduced from 3 to 2. The experimenter examined the staircase convergence after each session, and re-ran a session if she judged that the within-session staircases had not converged to approximately the same achromatic point. We subsequently implemented our session exclusion criteria (see methods for Experiment 1) and applied these to all sessions. As a result of the exclusion, there were some conditions where only a single session of data was available for an observer. In addition, there were no good sessions for Observer JOP for the matte disk test at 3 cd/m^2 for either the consistent-cue or reduced-cue conditions.

Results

Figure 5A shows the constancy indices for both test luminances from Experiment 2 for the case where the contextual spheres were matte (as they were in Experiment 1). There was no significant effect of test luminance, nor did test luminance interact with the effect of test object or background (Table 2). For the 5 cd/m^2 tests, the pattern

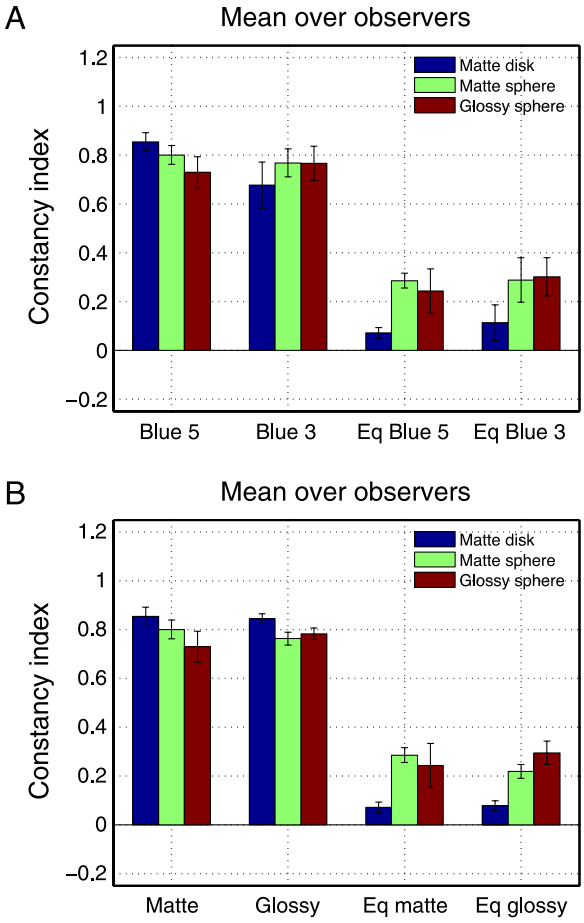

Figure 5. Constancy indices, Experiment 2. (A) Mean across observer constancy indices for Experiment 2 for conditions with matte contextual spheres for both test luminances. Within group, the three bars provide indices for the matte disk, matte sphere, and glossy sphere. Blue 5: Blue illuminant change, consistent-cue, 5 cd/m^2 test. Blue 3: Blue illuminant change, consistent-cue, 3 cd/m^2 test. Eq Blue 5: Blue illuminant change, reduced-cue, 5 cd/m^2 test; Eq Blue 3: Blue illuminant change, reduced-cue, 3 cd/m^2 test. (B) Mean constancy indices for Experiment 2 with 5 cd/m^2 tests for both matte and glossy contextual spheres. Within group, the three bars provide indices for the matte disk, matte sphere, and glossy sphere. Matte: Blue illuminant change, consistent-cue, matte contextual spheres. Glossy: Blue illuminant change, consistent-cue, glossy contextual spheres. Eq Matte: Blue illuminant change, reduced-cue, matte contextual spheres. Eq Glossy: Blue illuminant change, reduced-cue, glossy contextual spheres. Note that the data plotted for Matte and Eq Matte in (B) are the same data plotted for Blue 5 and Eq Blue 5 in (A). Error bars in both panels show ±1 *SEM*. The Supplementary materials provide tabulated data for the individual observers, achromatic chromaticity plots for each observer, and constancy index plots for each observer.

of results for the matte disk and matte sphere replicate what was found from Experiment 1 (see Figure 4). For the 3 cd/m^2 tests, however, constancy for the matte disk was not better than that for the spheres in the consistent-cue condition. In addition, for both test luminances, there was no improvement in constancy between the matte and glossy spheres. What is robust across the two experiments is the improvement in constancy for the reduced-cue conditions between the matte disk and the spheres, and the data for Experiment 2 taken as a whole reveal a significant interaction between test object shape and cue condition (Table 2).

Figure 5B shows the constancy indices for both matte and glossy contextual spheres from Experiment 2, where the tests were 5 cd/m^2. There was no significant effect of contextual sphere material, nor did contextual sphere material interact with the effect of test object or background (Table 3). The same interaction between matte disk and spheres is seen for the glossy contextual spheres as was observed for the matte contextual spheres. Note that the data for the matte contextual spheres in Figure 5B are re-plotted from Figure 5A and do not represent additional measurements.

Discussion

We measured the color constancy of three-dimensional objects viewed in three-dimensional scenes and compared this to constancy for a flat matte test embedded in the same scenes. We found that when we changed only the illuminant (consistent-cue conditions) constancy was very good for all of our test objects. In aggregate, however, constancy was slightly better for the flat matte disk than for the test spheres in the consistent-cue conditions (see Figures 4B, 5A, and 5B; constancy was better for disk except for 3 cd/m^2 test luminance in Figure 5A). Our consistent-cue conditions are similar to the manipulations studied in most experiments that assess constancy. In addition, we studied reduced-cue conditions. These silence local contrast as a cue to the illuminant change, and by studying them we separate the effect of local contrast from other processes that mediate constancy. As expected, constancy is much reduced in the reduced-cue conditions. Moreover, and most interesting, we found that constancy was systematically worse for the matte disk than for the test spheres (again Figures 4B, 5A, and 5B). Although the effects of object shape are not huge, they are easily visible. To provide a sense for the interested reader, the online supplement includes, for each observer for the 3 cd/m^2 test conditions of Experiment 2, images that render the equivalent illuminant chromaticities for each object shape and cue condition. Our other manipulations (matte versus glossy sphere, matte versus glossy contextual spheres, test luminance) had little if any effect.

Our data present a fundamental challenge for generalizing theories of constancy developed for two-dimensional scenes to explain constancy for three-dimensional viewing. This challenge arises because there is no obvious intrinsic

Source	VarType	SumSq	*df*	MeanSq	*F*	*p*
Observer	Random	0.39	4	0.1	7.09	<0.001
Test object	Fixed	0.1	2	0.05	3.74	<0.05
Cue condition	Fixed	4.4	1	4.4	323	<0.001
Target luminance	Fixed	0	1	0	0.08	n.s. (0.79)
Test object * Cue condition	Fixed	0.12	2	0.06	4.34	<0.05
Test object * Target luminance	Fixed	0.03	2	0.01	0.96	n.s. (0.39)
Cue condition * Target luminance	Fixed	0.03	1	0.03	2.02	n.s. (0.16)
Error	Random	0.6	44	0.01		
Total		5.64	57			

Table 2. ANOVA for Experiment 2, matte contextual spheres, both test luminances. Observers were coded as a random variable. The main effect of each variable and the two-way interactions between all variables except observer were modeled. Data are for the two test luminances and the matte contextual spheres.

component of two-dimensional models that can explain an interaction between object shape and cue condition. Thus, additional theory must be added before our current understanding of constancy for two-dimensional scenes can be applied to natural viewing. We return to this point below.

In Experiment 1, we found that constancy for a glossy test sphere was better than for a matte sphere, qualitatively consistent with earlier reports that adding specular highlights to a scene improves constancy (Yang & Maloney, 2001; Yang & Shevell, 2003). This effect did not replicate in Experiment 2, however. In addition, when we performed a manipulation more directly comparable to the earlier work by introducing glossy contextual spheres, we did not find improved constancy. We do not have an explanation for the difference in results between our current experiments and the earlier work.

There have been very few other studies of constancy for three-dimensional objects. Hedrich et al. (2009) used a memory-based paradigm and found that constancy with respect to spectral illumination changes for consistent-cue conditions was better for three-dimensional objects embedded in three-dimensional scenes than for two-dimensional objects embedded in two-dimensional scenes. de Almeida et al. (2010) carried out a similar comparison using a different experimental paradigm and reached a different conclusion: They found that constancy was similar across a switch between two-dimensional and three-dimensional viewing. Our data do not speak directly to the issues addressed by these papers, since in all of our conditions the test object was viewed in a three-dimensional scene. We do note that in our hands, the relative degree of constancy for two-dimensional versus three-dimensional tests depends on cue conditions. There may be other cases where non-geometric aspects of the stimulus manipulations interact with the shape of the test, which could explain why different laboratories reach different conclusions.

Recent work from our laboratory examined constancy of perceived lightness and glossiness for three-dimensional objects rendered under spatially complex illumination fields (Olkkonen & Brainard, 2010, 2011). This work does not compare constancy for two- and three-dimensional objects but does reveal striking interactions between object shape, illumination field, and perceived glossiness. While different in many details from the experiments presented here, these interactions have similar conceptual implications to those of the present results, namely, that it may not be straightforward to use results obtained using one object shape to predict constancy for other object shapes.

An additional point of contact in the literature is that interactions similar to the ones we found have been

Source	VarType	SumSq	*df*	MeanSq	*F*	*p*
Observer	Random	0.08	4	0.02	2.71	<0.05
Shape	Fixed	0.04	1	0.04	4.72	<0.05
Context	Fixed	0	1	0	0	n.s. (0.98)
Cue condition	Fixed	5.5	1	5.5	706	<0.001
Shape * Context	Fixed	0	1	0	0	n.s. (0.98)
Shape * Cue condition	Fixed	0.24	1	0.24	30.4	<0.001
Context * Cue condition	Fixed	0	1	0	0.01	n.s. (0.92)
Error	Random	0.38	49	0.01		
Total		6.1	59			

Table 3. ANOVA for Experiment 2, 5 cd/m^2 tests, both matte and glossy contextual spheres. Observers were coded as a random variable. The main effect of each variable and the two-way interactions between all variables except observer were modeled. Data are for the 5 cd/m^2 test luminance and both the matte and glossy contextual spheres. Note that the matte contextual sphere 5 cd/m^2 data are the same as analyzed in Table 2.

observed when a flat test is moved in depth relative to a background; in that case, consistent-cue constancy is best when the test is co-planar with the background, while the opposite was found for reduced-cue conditions (Werner, 2006; see also Radonjić, Todorović, & Gilchrist, 2010).

How might we think about the interaction between test object shape and the nature of the spectral changes in the contextual image? In natural scenes, the illuminant typically varies both with the position of an object in the scene and with the direction of the incident light (Debevec, 1998; Dror, Willsky, & Adelson, 2004; Koenderink, Pont, van Doorn, Kappers, & Todd, 2007). This means that that the light reflected from each location on an object depends on the three-dimensional orientation of the surface normal at that location, so that the visual processing required for constancy can vary across locations within a single object. Brainard and Maloney (2011) review a theory of how the visual system takes surface normal into account, based on studies of how the pose of matte planar surfaces affects their appearance (Bloj et al., 2004; Boyaci et al., 2004, 2003; Ripamonti et al., 2004). Whether similar ideas can be applied to understand how individual locations within a single object are perceived is an open question, as is the question of whether it is fruitful to conceive of object color appearance as resulting from an integration of appearance computed at individual locations across the object's surface (but see Ling, Vurro, & Hurlbert, 2008; Olkkonen, Hansen, & Gegenfurtner, 2008; Xiao & Brainard, 2008). Investigation of these questions strikes us as a promising direction to pursue as we seek to understand the interaction between shape and constancy.

A second important observation, developed most fully in the literature on lightness perception, is that the relative influence of different locations in a contextual image on the appearance of a test patch varies with the position and pose of that patch (Adelson, 2000; Gilchrist, 2006; Gilchrist et al., 1999). Given this, it may also be the case that variation in the shape of an object has a similar influence, and this could in turn produce an interaction between shape and constancy.

Acknowledgments

This work was supported by NIH RO1 EY10016 and P30 EY001583. We thank J. Andrews-Labenski, C. Broussard, and M. Suplick for technical assistance. A. Radonjić provided helpful comments on the manuscript.

Commercial relationships: none.
Corresponding author: David H. Brainard.
Email: brainard@psych.upenn.edu.
Address: Department of Psychology, University of Pennsylvania, 3270 Walnut St., Philadelphia, PA 19104, USA.

Footnote

[1]The Supplementary material is available at http://color.psych.upenn.edu/supplements/constancy3Dobjects/.

References

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.

Allred, S. R., & Brainard, D. H. (2009). Contrast, constancy, and measurements of perceived lightness under parametric manipulation of surface slant and surface reflectance. *Journal of the Optical Society of America A, 26,* 949–961.

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

Blake, A., & Bülthoff, H. (1990). Does the brain know the physics of specular reflection. *Nature, 343,* 165–168.

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

Bloj, M., Ripamonti, C., Mitha, K., Greenwald, S., Hauck, R., & Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision, 4*(9):6, 735–746, http://www.journalofvision.org/content/4/9/6, doi:10.1167/4.9.6. [PubMed] [Article]

Boyaci, H., Doerschner, K., & Maloney, L. T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision, 4*(9):1, 664–679, http://www.journalofvision.org/content/4/9/1, doi:10.1167/4.9.1. [PubMed] [Article]

Boyaci, H., Maloney, L. T., & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision, 3*(8):2, 541–553, http://www.journalofvision.org/content/3/8/2, doi:10.1167/3.8.2. [PubMed] [Article]

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

Brainard, D. H. (2004). Color constancy. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 948–961). Cambridge, MA: MIT Press.

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

Brainard, D. H., Longere, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B. (2006). Bayesian model of human color constancy. *Journal of Vision, 6*(11):10, 1267–1281, http://www.journalofvision.org/content/6/11/10, doi:10.1167/6.11.10. [PubMed] [Article]

Brainard, D. H., & Maloney, L. T. (2011). Surface color perception and equivalent illumination models. *Journal of Vision, 11*(5):1, 1–18, http://www.journalofvision.org/content/11/5/1, doi:10.1167/11.5.1. [PubMed] [Article]

Brainard, D. H., Pelli, D. G., & Robson, T. (2002). Display characterization. In J. P. Hornak (Ed.), *Encyclopedia of imaging science and technology* (pp. 72–188). New York: Wiley.

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

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

CIE (2004). *Colorimetry* (3rd ed., No. 15.2004). Vienna: Bureau Central de la CIE.

CIE (2007). *Fundamental chromaticity diagram with physiological axes—Parts 1 and 2* (Technical Report 170-1). Vienna: Central Bureau of the Commission Internationale de l' Éclairage.

de Almeida, V. M. N., Fiadeiro, P. T., & Nascimento, S. M. (2010). Effect of scene dimensionality on color constancy with real three-dimensional scenes and objects. *Perception, 39,* 770–779.

Debevec, P. (1998). *Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography*. Paper presented at SIGGRAPH 98, Orlando, FL (pp. 189–198).

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision, 4*(2):1, 57–81, http://www.journalofvision.org/content/4/2/1, doi:10.1167/4.2.1. [PubMed] [Article]

Doerschner, K., Boyaci, H., & Maloney, L. T. (2004). Human observers compensate for secondary illumination originating in nearby chromatic surfaces. *Journal of Vision, 4*(2):3, 92–105, http://www.journalofvision.org/content/4/2/3, doi:10.1167/4.2.3. [PubMed] [Article]

Doerschner, K., Boyaci, H., & Maloney, L. T. (2007). Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. *Vision Research, 47,* 3409–3423. [PubMed]

Dror, R. O., Willsky, A. S., & Adelson, E. H. (2004). Statistical characterization of real-world illumination. *Journal of Vision, 4*(9):11, 821–837, http://www.journalofvision.org/content/4/9/11, doi:10.1167/4.9.11. [PubMed] [Article]

Fleming, R. W., & Bülthoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception, 2,* 346–382.

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5):3, 347–368, http://www.journalofvision.org/content/3/5/3, doi:10.1167/3.5.3. [PubMed] [Article]

Foster, D. H. (2011). Color constancy. *Vision Research, 51,* 674–700.

Funt, B. V., & Drew, M. S. (1993). Color space analysis of mutual illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15,* 1319–1326.

Gilchrist, A. (2006). *Seeing black and white*. Oxford, UK: Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V., & Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review, 106,* 795–834.

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

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics, 28,* 527–538.

Granzier, J. J. M., Brenner, E., Cornelissen, F. W., & Smeets, J. B. J. (2005). Luminance–color correlation is not used to estimate the color of the illumination. *Journal of Vision, 5*(1):2, 20–27, http://www.journalofvision.org/content/5/1/2, doi:10.1167/5.1.2. [PubMed] [Article]

Griffin, L. D. (1999). Partitive mixing of images: A tool for investigating pictorial perception. *Journal of the Optical Society of America A, 16,* 2825–2835.

Hedrich, M., Bloj, M., & Ruppertsberg, A. I. (2009). Color constancy improves for real 3D objects. *Journal of Vision, 9*(4):16, 1–16, http://www.journalofvision.org/content/9/4/16, doi:10.1167/9.4.16. [PubMed] [Article]

Helson, H., & Jeffers, V. B. (1940). Fundamental problems in color vision: II. Hue, lightness, and saturation of selective samples in chromatic illumination. *Journal of Experimental Psychology, 26,* 1–27.

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

Hochberg, J. E., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology, 47,* 263–266.

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

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

Kim, J., Marlow, P., & Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision, 11*(9):4, 1–19, http://www.journalofvision.org/content/11/9/4, doi:10.1167/11.9.4. [PubMed] [Article]

Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M., & Todd, J. T. (2007). The visual light field. *Perception, 36,* 1595–1610. [PubMed]

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 307–312.

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

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

Ling, Y., Vurro, M., & Hurlbert, A. (2008). *Surface chromaticity distributions of natural objects under changing illumination*. Paper presented at 4th European Conference on Colour in Graphics, Imaging and Vision (pp. 263–267).

Marlow, P., Kim, J., & Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *Journal of Vision, 11*(9):16, 1–12, http://www.journalofvision.org/content/11/9/16, doi:10.1167/11.9.16. [PubMed] [Article]

McCann, J. J., McKee, S. P., & Taylor, T. H. (1976). Quantitative studies in retinex theory: A comparison between theoretical predictions and observer responses to the 'Color Mondrian' experiments. *Vision Research, 16,* 445–458.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447,* 206–209.

Newhall, S. M., Nickerson, D., & Judd, D. B. (1943). Final report of the O.S.A. subcommittee on the spacing of Munsell Colors. *Journal of the Optical Society of America, 33,* 385–412.

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

Nishida, S., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A, 15,* 2951–2965.

Olkkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision, 10*(9):5, 1–19, http://www.journalofvision.org/content/10/9/5, doi:10.1167/10.9.5. [PubMed] [Article]

Olkkonen, M., & Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception, 2,* 1014–1034. [Article]

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2008). Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision, 8*(5):13, 1–16, http://www.journalofvision.org/content/8/5/13, doi:10.1167/8.5.13. [PubMed] [Article]

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *Journal of Vision, 9*(12):6, 1–18, http://www.journalofvision.org/content/9/12/6, doi:10.1167/9.12.6. [PubMed] [Article]

Pessoa, L., Mingolla, E., & Arend, L. E. (1996). The perception of lightness in 3-D curved objects. *Perception & Psychophysics, 58,* 1293–1305.

Radonjić, A., Todorović, D., & Gilchrist, A. (2010). Adjacency and surroundedness in the depth effect on lightness. *Journal of Vision, 10*(9):12, 1–16, http://www.journalofvision.org/content/10/9/12, doi:10.1167/10.9.12. [PubMed] [Article]

Ripamonti, C., Bloj, M., Hauck, R., Mitha, K., Greenwald, S., Maloney, S. I., & Brainard, D. H. (2004). Measurements of the effect of surface slant on perceived lightness. *Journal of Vision, 4*(9):7, 747–763, http://www.journalofvision.org/content/4/9/7, doi:10.1167/4.9.7. [PubMed] [Article]

Sharon, L., Li, Y., Motoyoshi, I., Nishida, S., & Adelson, E. H. (2008). Image statistics for surface reflectance perception. *Journal of the Optical Society of America A, 25,* 846–865.

Stockman, A., & Sharpe, L. T. (2000). Spectral sensitivities of the middle- and long-wavelength sensitive

cones derived from measurements in observers of known genotype. *Vision Research, 40,* 1711–1737.

Todd, J. T., Norman, J. F., & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science, 15,* 33–39.

Tominaga, S., & Wandell, B. A. (1989). The standard surface reflectance model and illuminant estimation. *Journal of the Optical Society of America A, 6,* 576–584.

Vos, J. J. (1978). Colorimetric and photometric properties of a 2 degree fundamental observer. *Color Research and Application, 3,* 125–128.

Ward, G. J. (1992). *Measuring and modeling anisotropic reflection*. Paper presented at SIGGRAPH 92: Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques (pp. 265–272).

Ward, G. J. (1994, July 24–29). *The RADIANCE lighting simulation and rendering system*. Paper presented at Proceedings of ACM SIGGRAPH 94 (pp. 459–472).

Werner, A. (2006). The influence of depth segregation on colour constancy. *Perception, 35,* 1171–1184.

Xiao, B., & Brainard, D. H. (2008). Surface gloss and color perception of 3D objects. *Visual Neuroscience, 25,* 371–385.

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

Yang, J. N., & Shevell, S. K. (2003). Surface color perception under two illuminants: The second illuminant reduces color constancy. *Journal of Vision, 3*(5):4, 369–379, http://www.journalofvision.org/content/3/5/4, doi:10.1167/3.5.4. [PubMed] [Article]