Cues to an Equivalent Lighting Model

Huseyin Boyaci   Department of Psychology, University of Minnesota, Minneapolis, MN, USA

Katja Doerschner   Department of Psychology, New York University, New York, NY, USA

Laurence T. Maloney   Department of Psychology and Center for Neural Science, New York University, New York, NY, USA

We investigate how human observers make use of three candidate cues in their lightness judgments. Each cue potentially provides information about the spatial distribution of light sources in complex, rendered 3D scenes. The illumination (lighting model) of each scene consisted of a punctate light source combined with a diffuse light source. The cues were (1) cast shadows, (2) surface shading, and (3) specular highlights. Observers were asked to judge the albedo of a matte grayscale test patch that varied in orientation with respect to the punctate light source. We tested their performance in scenes containing only one type of cue and in scenes where all three cue types were present. From the results, we deduced how accurately they had estimated the spatial distribution of light sources in each scene given the cues available. In Experiment 1, we established that all of the individual cues were used in isolation. We showed that the highlight and cast shadow cues in isolation were used by more than half of the observers. We could reject the hypothesis that the observers did not make use of the shading cue for only one observer. In Experiment 2, we showed that the observers combined information from multiple cues when all three cues were presented together.

Keywords: surface color perception, albedo, cue combination, lightness

Introduction

In everyday scenes, the light that reaches the eye from a matte surface patch depends on many factors. These include the surface properties of the patch and also the lighting model (the spatial distribution of light sources in the scene). When the spatial distribution of the light sources is not uniform, the light reflected by a matte surface can also depend on its orientation (Figure 1).

Boyaci, Maloney, and Hersh (2003) and Ripamonti et al. (2004) demonstrated that human observers can take the spatial distribution of light sources into account in their lightness estimates, suggesting that the visual system develops an estimate of some aspects of the lighting model. As early as the study of von Helmholtz (1866/1924), researchers considered how the lightness or color of a surface could be estimated by first inferring the lighting model from the cues present in the scene. Typically, however, these discussions centered on estimation of the chromaticity of a single light source illuminating the scene, and there is no previous study demonstrating that, in arriving at a lightness estimate, the visual system uses any specific cue for estimating the spatial variation of light falling on the test surface from all directions. In this article, we examine three candidate cues that may affect the observers' lightness estimates.

Evidence that human observers develop an estimate of the lighting model

Boyaci et al. (2003) studied how changes in test patch orientation affected perception of the albedo of achromatic surfaces in rendered scenes. The lighting model used in rendering the scenes consisted of a punctate source and a diffuse source, both neutral in chromaticity. The punctate source was placed behind the observers and was not directly visible to them. Boyaci et al. varied test patch orientation from trial to trial to examine how changes in orientation affected the observers' perception of the surface albedo of the test patch. They found that the observers partially discounted the effect of changing orientation in judging surface albedo. Ripamonti et al. (2004) repeated this experiment using actual surfaces and lights with the same outcome (see also Bloj et al., 2004).

Boyaci, Doerschner, and Maloney (2004) examined the observers' performance in an achromatic setting task with a lighting model consisting of a yellow punctate source and a diffuse blue source. They found that the observers partially discounted the effect of changing orientation in judging surface color. For each of the experiments just described, it was possible to estimate the perceived location of the punctate light source directly from the observers' responses.

Figure 1. Orientation changes and Lambert's law. Lambert's law (Haralick & Shapiro, 1993) is a mathematical description of absorption and emission of light by an idealized achromatic matte surface. Identical square surfaces with different orientations and locations were rendered with a distant punctate light source placed along the line of sight, perpendicular to the figure. It is easy to see the large differences in emitted luminance with changes in orientation.

The combined results described above imply that the observer's visual system effectively develops an estimate of the spatial and spectral distribution of parts of the lighting model that are not directly visible. We refer to this estimate as the observer's Equivalent Lighting Model (ELM).

Boyaci et al. (2003) proposed that the observers go about the task of estimating the surface albedo of an achromatic surface in two stages (Figure 2). First, they estimate the parameters of the lighting model: the total light intensity $\hat{E}$, the direction to the punctate light source $\hat{\Upsilon}_P = (\hat{\psi}_P, \hat{\phi}_P)$ (azimuth and elevation), and the ratio of the punctate light to total light $\hat{\pi}$ illuminating the scene. The resulting ELM is denoted as $\hat{\mathbf{M}} = [\hat{E}, \hat{\Upsilon}_P, \hat{\pi}]$. Observers also estimate the surface orientation $\hat{\Upsilon}_T$ of the test surface. In Stage 2, they estimate surface albedo, $\hat{\alpha}$, by computing a geometric correction factor (explained below) based on $\hat{\Upsilon}_T$ and the possibly erroneous estimates of the ELM from Stage 1. Note that the perceived luminance, $\hat{L}$, is assumed to be readily available to the visual system.

Cues to the lighting model

The primary question we address here is, what sources of information in the scene does the visual system use to develop an estimate of ELM?

Specifically, we examine what sources of information the visual system uses to estimate the location of neutral punctate light sources and their relative intensities in a scene. We refer to these sources as "cues to the lighting model," and in this article we will be concerned with the cues to the location of the punctate light source and the punctate-total ratio in the lightness matching task of Boyaci et al. (2003).

There are three candidate cues that we will consider (see Figures 4 and 8): cast shadows, surface shading (attached
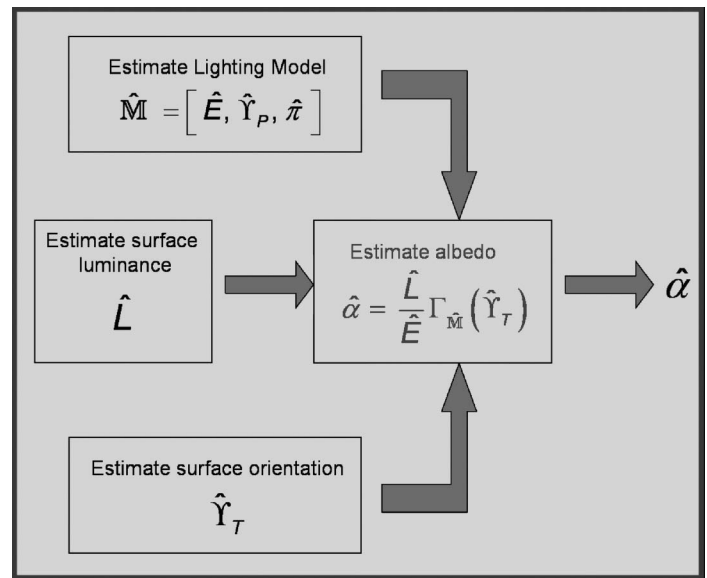

Figure 2. A two-stage model of albedo estimation in human vision. The visual system combines estimates of the lighting model and test surface orientation to compute the geometric correction function for a particular scene. Errors in these estimates will lead to patterned errors in perception of albedo.

shadows), and the virtual image formed in the surface of a specular object. We refer to this last cue as "specular highlights" or "highlights." These three cues were always present in the scenes used by Boyaci et al. (2003, 2004) and two of the cues were present in the scenes of Ripamonti et al. (2004; there were no specular objects). Here we wish to determine, primarily, which of the three serve as cues in estimating $\hat{\Upsilon}_P$ and π. If we find that multiple cues are utilized by the visual system in a scene, we wish to test whether the visual system combines them to arrive at less variable estimates of the parameters of the lighting model (Ernst & Banks, 2002; Landy, Maloney, Johnston, & Young, 1995; Oruç, Maloney, & Landy, 2003).

Two previous studies have examined how the visual system uses cues present in the scenes to arrive at estimates of the chromaticity of the illuminant (Kraft & Brainard, 1999; Yang & Maloney, 2001; see also Maloney, 2002). Snyder, Doerschner, and Maloney (2005) showed that inclusion of glossy spherical objects improved the observers' lightness constancy in computer-rendered 3D scenes. Koenderink, van Doorn, and Pont (2004) studied how human observers estimate light source direction in simple scenes using information from cast shadows and shading. Gilchrist (1977, 1980) demonstrated that errors in estimating the lighting model could result in large errors in estimating albedo and Gilchrist et al. (1999) review possible heuristics for estimating lighting models in simple scenes.

Similar problems are considered in computer vision and the computer graphics literature. Hara, Nishino, and Ikeuchi (2005) developed computer algorithms using specular highlights and shading of matte surfaces to estimate the light source position. Wang and Samaras (2003) were able to
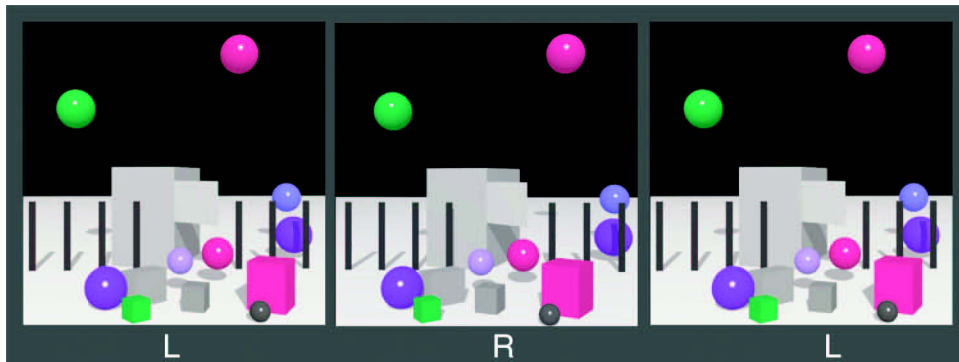
Figure 3. Sample stereo pair. The left pair is for uncrossed fusion, the right pair is for crossed fusion.

estimate lighting model using cast shadows and shading first independently and then by integrating the two results. Li, Lin, Lu, and Shum (2003) developed an algorithm that could determine distribution of multiple light sources by integrating information from shading, shadows, and specular reflections.

Experiment 1

Methods
Stimuli

The stimuli were computer-rendered, three-dimensional scenes composed of simple geometrical objects such as spheres and cubes with different colors and reflectance properties. These objects were intended to serve as candidate cues to provide information about the lighting model. Each scene was rendered twice with slightly different viewpoints corresponding to the positions of the observers' eyes. A sample stereo pair is shown in Figure 3.

The scenes were illuminated by a mixture of diffuse and punctate light sources. A gray matte (Lambertian) test patch was placed near the center of the scene whose orientation and luminance varied from trial to trial as explained below. The scenes differed in the kinds of cues they contained.

There were a total of four cue conditions. In the cast shadow condition, all objects but the flat[1] matte poles and floor were removed (Figure 4A). In the shading condition, everything except the Lambertian cubic objects was removed (Figure 4B). In the highlights condition only the glossy spheres were present (Figure 4C). In the all cues condition, we presented the scene with all cues available (Figure 4D).

Software and apparatus

The experimental code was written by us in the C language using the X Window System, Version 11R6 (Scheifler & Gettys, 1996). The scenes were rendered by the open source physics-based rendering package Radiance (Larson & Shakespeare, 1996). The Radiance output for each image consisted of floating point triplets for each pixel. These triplets were interpreted as the relative retinal illumination if the observers' eyes were at the viewpoints selected in the virtual scene. We translated these relative luminance values into 24-bit graphics codes after correcting them for nonlinearities in the monitor responses by means of measured look-up tables.

The rendered images were presented to the observers binocularly in a stereoscope, on two 21-in. Sony Trinitron Multiscan GDM-F500 monitors, both having a highest luminance of 115 cd/m^2. The computer used in the experimental apparatus was a Dell Workstation with an Nvidia
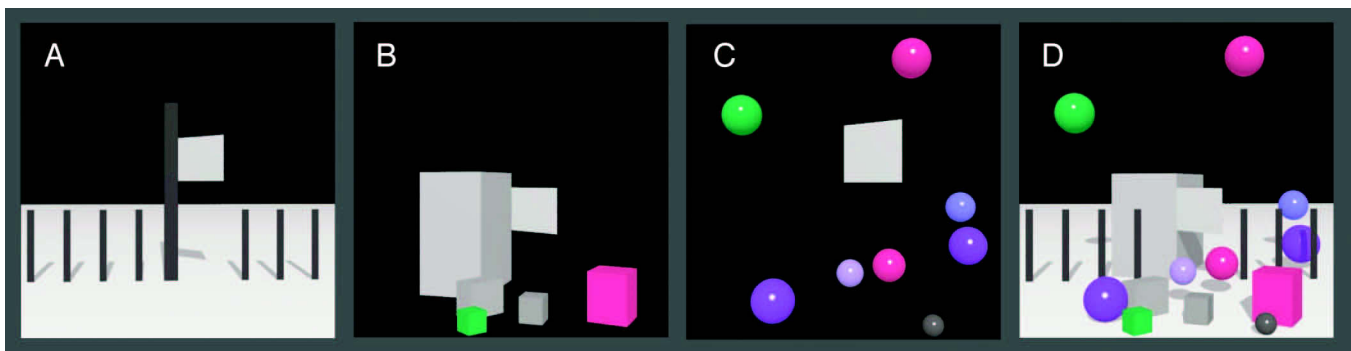

Figure 4. Cue conditions. We consider three cues that potentially could be used to estimate the lighting model in the scene: cast shadows, shading of matte surfaces, and specular highlights. (A) Cast shadows condition. (B) Shading condition. (C) Highlights condition. (D) All cues condition.

dual-head graphics card that controlled both monitors. The experimental program ran under Fedora Core Linux, Version 3 (for details, see Boyaci et al., 2003).

Test patch

The test patch was placed near the center of the scene, 70 cm away from the observers' viewpoint. The same distance of 70 cm was also the physical distance from the observer's eyes to the computer screens. The test patch could appear at any one of the five different orientations specified by the azimuth ψ_T of the normal to the test patch surface (elevation $\varphi_T = 0$ remained constant). The values ψ_T could take on were −60, −45, 0, 45, and 60 deg. The luminance of each test patch was kept constant independent of its orientation. Four different luminance values were interleaved randomly to avoid possible artifacts. These luminance values were 0.44, 0.50, 0.56, and 0.61 proportion of the maximum possible luminance for the computer monitors (115 cd/m^2).

Light sources

The scenes were illuminated by a neutral punctate and a neutral diffuse light source. The punctate source was placed directly above and behind the observer's head.[2] The distance between the test patch and the punctate source was 670 cm (580 cm front, 335 cm above). The actual direction to the punctate source was $\Upsilon_P = (\psi_P, \varphi_P) = (0, 30$ deg). The position of the punctate source was never varied during the experiment and its location was sufficiently far from the test patch so that it could be treated as a collimated source across the extent of the test patch. The actual punctate-total light ratio was $\pi = 0.67$.

Task

The task of the observer was to indicate the lightness of the test patch by choosing a matching chip among a randomly ordered scale of chips presented near the right edge of the screen. The exact phrase used was "choose the matching chip which looks like it is cut out of the same piece of gray paper as the test patch." The scale was presented monocularly, to the right eye, and its ordering was changed randomly at each trial. The luminance values of the chips were 0.11, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.85, and 0.97 proportion of maximum luminance of the monitors (115 cd/m^2).

Procedure

Observers repeated each of 80 conditions 10 times (4 cue conditions, 4 luminances, 5 orientations). Between each trial, a uniform gray mask covering the entire image area was presented for 1 s. Different cue conditions were blocked in different sessions. Observers were allowed to perform a few trials before the actual experiment started. There were no time constraints. Each session took about half an hour to complete on the average.

Observers

Six observers completed the experiment. All were undergraduate students in New York University who were unaware of the purpose of the experiment.

Analysis and results

Parameters of the ELM

What would a lightness constant observer do and what are his or her possible errors? The albedo of a Lambertian surface under the lighting model $\mathbf{M} = [E, \Upsilon_P, \pi]$ is given by

$$\alpha = \frac{L}{E}\Gamma_{\mathbf{M}}(\Upsilon_T), \tag{1}$$

where L is its luminance and $\Gamma_{\mathbf{M}}(\Upsilon_T)$ is a geometric correction function,

$$\Gamma_{\mathbf{M}}(\Upsilon_T) = [\pi \cos \theta(\Upsilon_T, \Upsilon_P) + (1-\pi)]^{-1}, \tag{2}$$

where $\theta(\Upsilon_T, \Upsilon_P)$ is the angle between the normal to the surface and direction to the punctate light source. Intuitively, the geometric correction function is the amount that the "ideal albedo observer" must "correct" perceived albedo to compensate for the effect of orientation. Furthermore, $\cos\theta$ is given by

$$\cos \theta(\Upsilon_T, \Upsilon_P) = \cos \varphi_P \cos(\psi_T - \psi_P). \tag{3}$$

Employing Equation 1, one can show that the ratio of the luminance of the matching chip and the test patch is given by the *setting ratio*

$$\Lambda = \frac{L^R}{L^T} = C\Gamma_{\mathbf{M}}, \tag{4}$$

when their albedos match. In Equation 4, C characterizes the combined effect of the unknown absolute total light E in the scene and the unknown lighting model of the matching chip (for more details, see Boyaci et al., 2003).

We can depict any observer's geometric correction function as his setting ratios Λ plotted versus ψ_T. If the lightness constant Lambertian observer's visual system had perfect estimates of Υ_P and π, then his or her geometric correction function should have the same shape as the actual geometric
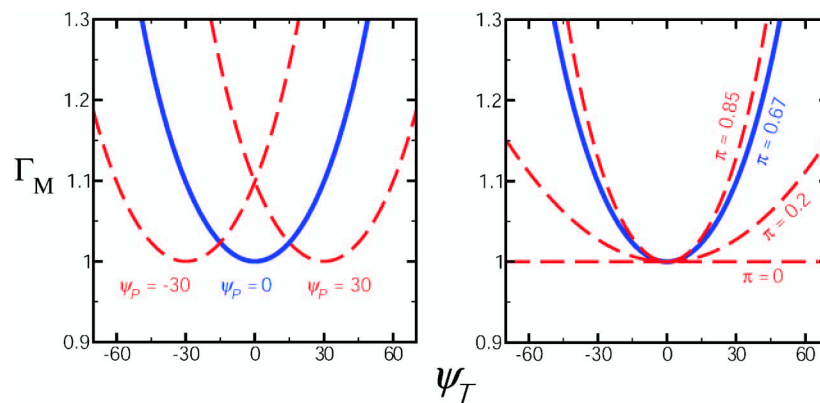
Figure 5. The effect of varying the lighting model parameters ψ_P and π on the geometric correction function. (Left) Γ_M versus ψ_T for different values of ψ_P (with $\varphi_T = \varphi_P = 0$). The curve in blue corresponds to the case $\psi_P = 0$ deg. The two dashed curves correspond to the cases $\psi_P = \pm30$ deg. Each curve represents the decrease in emitted luminance as a surface of constant albedo is rotated away from the punctate source. Changing the elevation φ_T of the surface normal leads to a similar set of curves. These are plotted in Boyaci et al. (2004). (Right) Γ_M versus ψ_T for various values of the punctate-total ratio π (with $\psi_P = 0$ deg). The minimum of Γ_M is always at $\psi_T = \psi_P$, the direction to the light; changing π affects only curvature. The blue, solid curve corresponds to $\pi = 0.67$ (the value we use in Experiment 1). The red, dashed curves correspond to values of $\pi = 0.87$ (more curved) and $\pi = 0.2$ (less curved). When $\pi = 0$, the curve becomes a horizontal line ($\Gamma_M \equiv 1$), a readily interpretable result: if $\pi = 0$ then the light is perfectly diffuse and changes in orientation do not affect the intensity of light absorbed by or emitted by the test surface.

function, up to an unknown scaling factor C. Conversely, given the form of the observer's geometric correction function, we can compute the estimates of $\hat{\psi}_P$ and $\hat{\pi}$, which accounts best for the data: misestimates of light azimuth, ψ_P, shift the function to the right or to the left. Misestimates of punctate-total balance, π, change the curvature of the function (Figure 5). As explained in Appendix B, all analyses were carried out using a transformation of the parameter π. We retain π in the text because it is readily interpreted as the punctate-total ratio defined above.

Results

We present the data of one observer (SRH) in Figure 6. We plot the average of the setting ratios, Λ, for all four test patches of different luminances at each orientation. The solid line is the best fit obtained by a maximum like-

lihood estimation analysis. We fit Equation 4 to the observer's data by the method of maximum likelihood (Mood, Graybill, & Boes, 1974, pp. 276 ff) under the assumption that the observer's settings are perturbed by Gaussian error with mean 0 and variance σ^2.

We tested the null hypothesis that $\pi = 0$ versus the alternative that $\pi \neq 0$ by a nested hypothesis test (Mood et al., 1974, pp. 440–442) as follows. We first obtain maximum likelihood estimates of the parameters $\hat{\psi}_P, \hat{\pi}, \hat{C}$, and $\hat{\sigma}^2$ without constraining $\hat{\pi}$. These are estimates of the parameters with the unconstrained model. We then obtain estimates of C and σ^2 under the constraint $\pi = 0$ (the null hypothesis) and without fitting the parameter ψ_P (in the $\pi = 0$ case, the observer's geometric correction function is a horizontal line, hence there is no minimum, or in other words the azimuth of the punctate source is not defined because there is no punctate source).
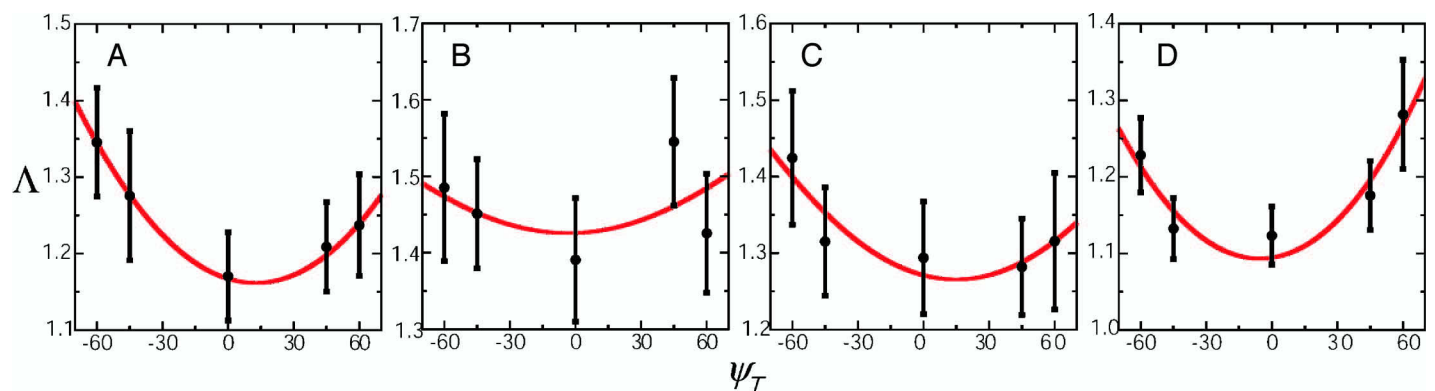

Figure 6. Experiment 1: Results of observer SRH under all four cue conditions. (A) Cast shadows condition, (B) shading condition, (C) highlights condition, and (D) all cues condition.
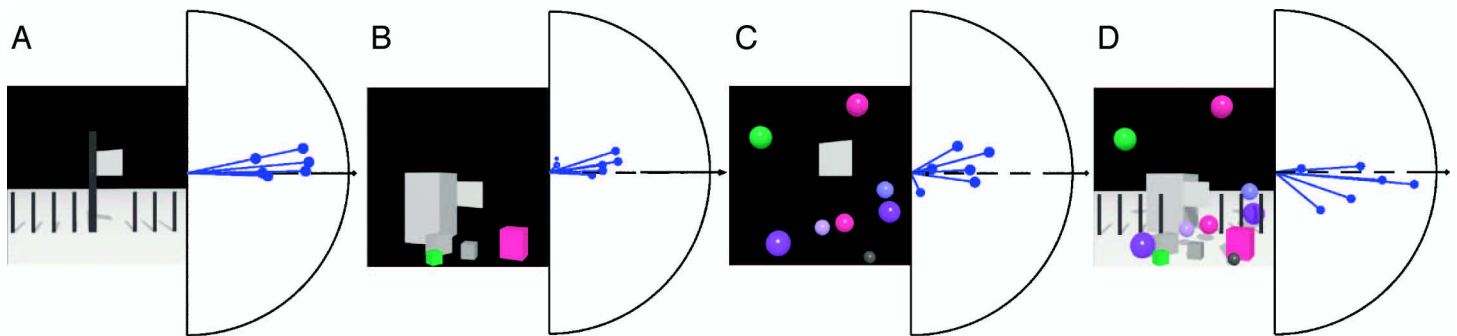
Figure 7. Experiment 1: Estimates of ψ_P and π plotted in polar coordinates. (A) Cast shadows condition, (B) shading condition, (C) highlights condition, and (D) all cues condition. The results for each observer are plotted as a blue line segment with one end at the origin. Each line segment has length π and orientation ψ_P. The true value of ψ_P is plotted as a dashed black line segment. It can be seen that the observers' estimates ψ_P are close to the true value in all cue conditions. The true value of the punctate-total ratio π is 0.67. It would lie outside each of the plots (radius corresponds to $\pi = 0.3$). Observers markedly underestimated π.

If λ denotes the ratio of the maximum likelihood achieved in fitting the unconstrained and the constrained models, then $-2 \log \lambda$ is approximately distributed as a χ_2^2 distribution under the null hypothesis (Mood et al., 1974, pp. 440–442). For a test of size p, we reject the null hypothesis $\pi = 0$ if and only if $-2 \log \lambda$ is greater than the $1 - p$ quantile of the χ_2^2 distribution.

Figure 7 summarizes the estimates of $\hat{\psi}_P$ and $\hat{\pi}$ of all observers in each of the four conditions. The estimates for each observer are plotted in polar coordinates (following Ripamonti et al., 2004) as blue line segments with one end at the origin. The angle between the segment and the horizontal is $\hat{\psi}_P$ and the distance from the origin is $\hat{\pi}$. A blue dot marks the end of the line segment. The true value of ψ_P is marked by a black dotted line segment. The true value $\pi = 0.67$ lies outside the plot.

Observers markedly underestimated the true value π in all conditions in agreement with the results in previous experiments with rendered scenes (Boyaci et al., 2003, 2004) and real scenes (Ripamonti et al., 2004). It can be seen that the estimates of ψ_P cluster around the true value in all conditions. Observers perceive the direction to the hidden punctate light source (nearly) accurately but underestimate its intensity relative to the diffuse source.

In Table 1, we report the exact p values of the nested hypothesis test described above for each condition and subject. If, in any one cue condition, the observers made no use of the

cues provided, then these p values would be distributed as uniform random variables in the unit interval [0,1]. We use this fact to test the hypothesis that none of the observers is using a specific cue by computing the logarithm of the geometric mean of the p values ($\log_{10} \overline{p}$) for that cue across the six subjects and comparing the result to the distribution of the geometric mean of six independent uniform (0,1) random variables (see Appendix A). We report the p values for this omnibus test statistic in the last column of Table 1.

These results indicate that at least some of the observers are making use of each of the individual cues and the combined cues. Examination of the p values for individual subjects suggests that one subject (SB) made use of no cues and we cannot claim that he or she is correcting albedo estimates for orientation in any of the conditions. It is plausible that some subjects are making use of at least two of the distinct cues, which we further examine in the next experiment.

Experiment 2

Methods
Stimuli

The stimuli were computer-rendered 3D scenes similar to the ones used in Experiment 1. We investigate the same

Cue condition	Observers						Overall
	MS	CD	SH	SRH	KP	SB	
All cues	.023	.0002	.002	.000003	.002	.58	<.0001
Cast shadows	.006	.0072	.02	.0002	.08	.078	<.0001
Shading	.02	.0572	.26	.5	.233	.78	.0495
Highlights	.24	.13	.001	.022	.78	.44	.0023

Table 1. Exact p values corresponding to the nested hypothesis test of the null hypothesis $\pi = 0$. The last column contains an omnibus p value corresponding to the hypothesis that no observer made use of the cues available in a given cue condition. Values greater than .05 are shaded (null hypothesis, $\pi = 0$, is not rejected).
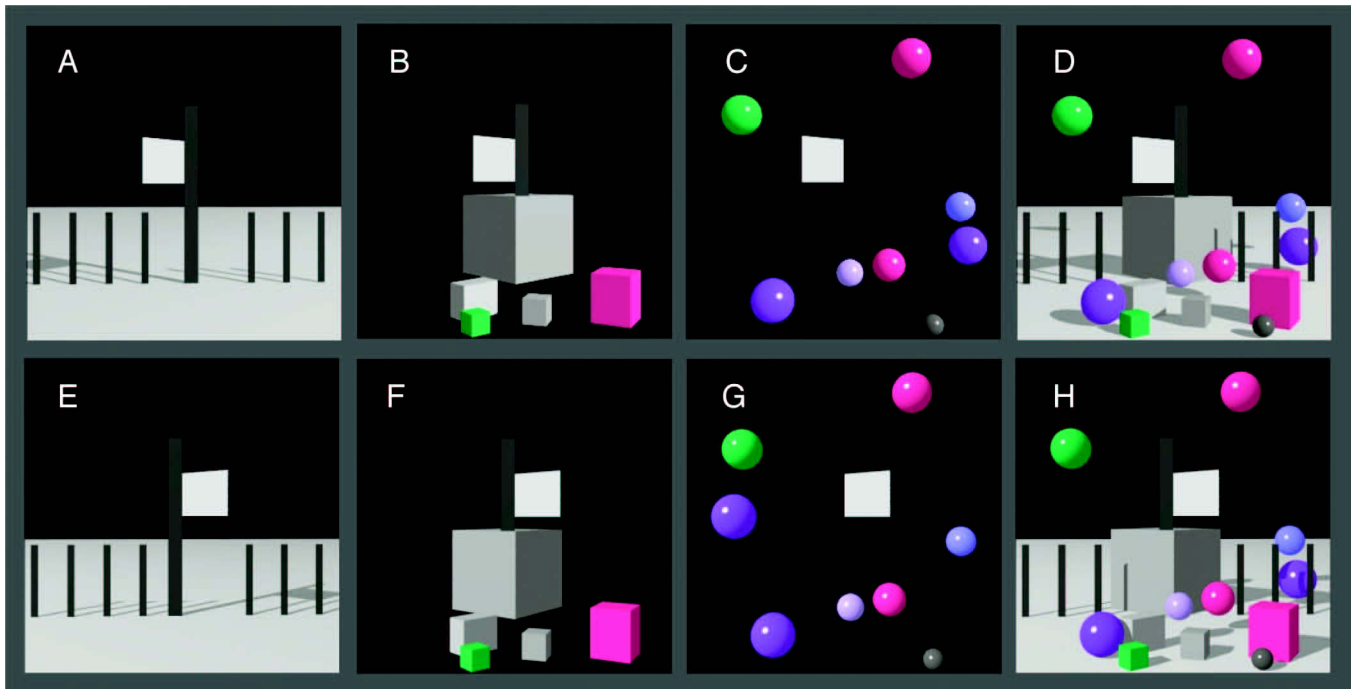
Figure 8. Experiment 2: Stimuli. The stimuli used in Experiment 2 were similar to those used in Experiment 1, except that there were two light conditions in Experiment 2. (A–D) Light from the right; (E–H) light from the left.

four cue conditions but with two different light positions as shown in Figure 8.

Software and apparatus

Same as in Experiment 1.

Test patch

Test patches had similar characteristics as in Experiment 1, except their luminances were 0.604, 0.678, 0.752, and 0.83 proportion of the maximum possible luminance of the computer monitors (115 cd/m^2). They could take on the same orientations $\psi_T = \{-60, -45, 0, 45, 60 \text{ deg}\}$.

Light sources

The scenes were illuminated by a combination of neutral punctate and diffuse light sources. The punctate source was placed above, behind, and either on the left or on the right of the observer's head. The distance between the test patch and the punctate source was 668 cm (409 cm front, 335 cm above, and 409 cm on the right or on the left). The direction to the punctate source was $\Upsilon_P = (\theta_P, \varphi_P) = (\pm 45, 30 \text{ deg})$. The position of the punctate source did not vary within a block. Punctate-total ratio was $\pi = 0.85$.

Task

Same as in Experiment 1.

Procedure

Observers repeated each of 160 conditions 10 times (2 light positions, 4 cue conditions, 4 luminances, and 5 orientations).

Observers

A total of 20 observers participated in a preliminary test session and two observers, one naïve (IB) and one of the authors (KD), completed the entire experiment. We set a rather restrictive threshold to decide on eligibility: Observers first ran the two all cues conditions (one in light on the left condition, other in light on the right condition) and only those whose lightness estimates showed a very significant correction ($p < .0005$) against changes in orientation were chosen (it is worth noting that nearly all 20 observers corrected their lightness settings for the test patch orientation in the all cues conditions at $p = .05$ level). An analysis of cue combination is more powerful for observers who correct their lightness estimates more in the all cues condition. For instance, data from an observer like SB in Experiment 1 would not be suitable to study cue combination because that observer did not use any of the cues. Only subjects who met criterion ran in the remaining experimental (single cue) conditions.

Analysis and results

We present the data of observer IB in Figure 9 in a similar fashion as in Experiment 1. We repeated the same
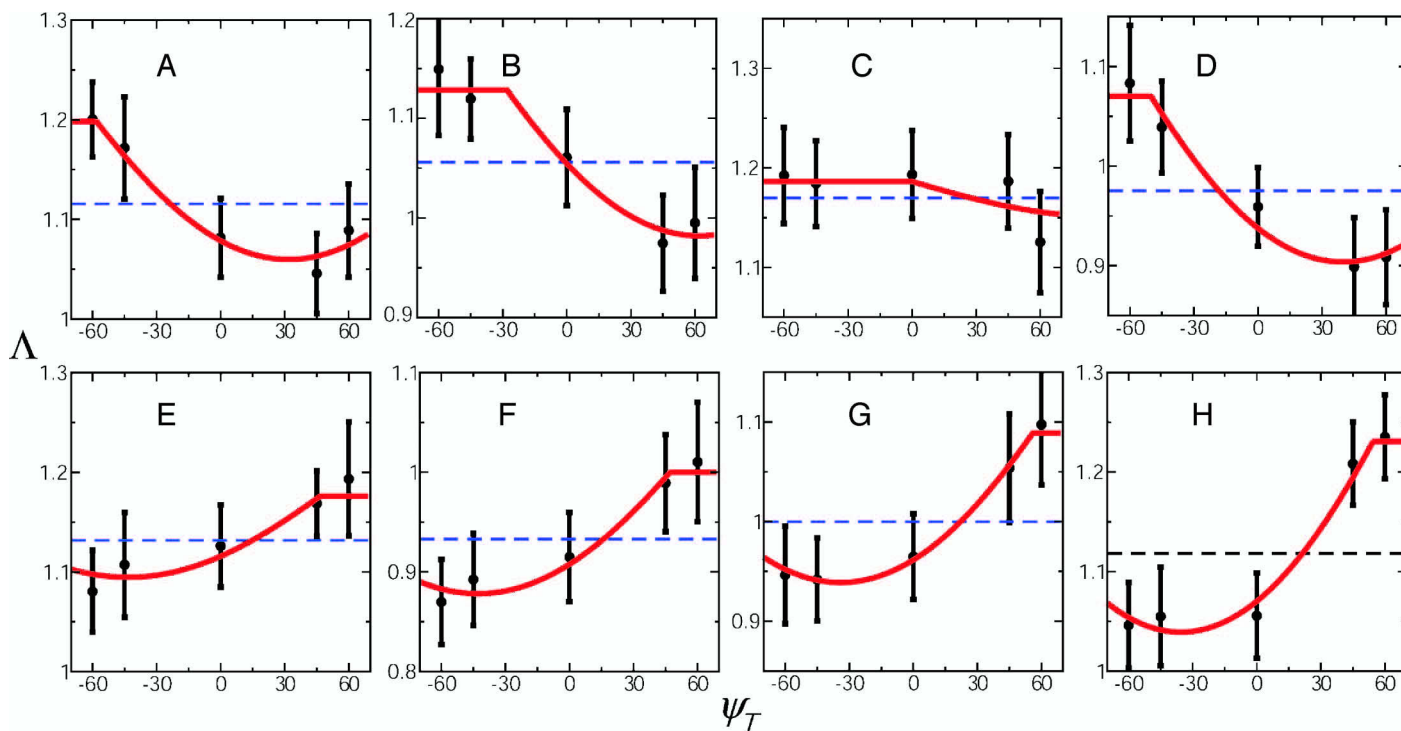
Figure 9. Experiment 2: Results (observer IB). Same as in Figure 6, except (A–D) light from the right; (E–H) light from the left.

statistical analysis described in the Analysis and results section of Experiment 1. Figure 10 summarizes the estimates of light position and punctate-total ratio (ψ_P, π) for both observers.

In all conditions (two light, four cue, total of eight), the null hypothesis that $\pi = 0$ (no correction) is rejected except for IB in the highlights condition under the light from the right and for KD in the highlights condition under the light from the right. We conclude that each of the individual cues and combined cues is used by all observers.

Given that the observers make use of each of the individual cues, we next consider whether any of the observers make use of more than one cue in the all cues condition.

Specifically, do they combine two or more cues to obtain an estimate of ψ_P and π that is more reliable than the estimates based on the cues in isolation?

Are cues combined to produce more reliable estimates?

Suppose that we have independent biased (or unbiased) Gaussian estimates from each one of the multiple cues

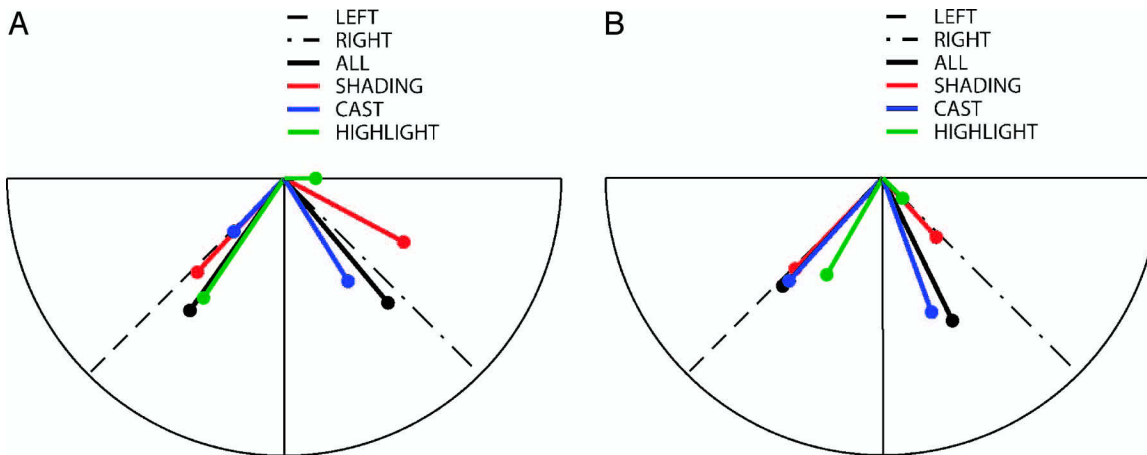$$\hat{\pi}_i \sim \Phi(B, \sigma_i^2) \quad i = 1, 2, 3, \tag{5}$$

Figure 10. Experiment 2: Results, both observers. (A) Observer IB, (B) observer KD. Similar to Figure 7, except all eight conditions (two light position, four cue conditions) are plotted in a single plot for each observer. Radius corresponds to $\pi = 0.3$. The true value $\pi = 0.85$ lies outside the plots.

Cue condition	Observers			
	KD—L	KD—R	IB—L	IB—R
All cues	33.8	19.1	53.3	36.1
$\hat{r}_{max}$	25.1	13.4	26.6	31.4
$\hat{r}_0$—Predicted optimum	59.3	23.1	57.5	51.1
Cast shadows	21.7	9.6	12.9	19.7
Shading	25.1	13.4	26.6	31.4
Highlights	12.5	0.02	18	0.02
z score (1)	−4.01	−5.10	−8.38	−2.00
z score (2)	10.04	3.30	1.25	6.08

Table 2. Computed and empirical estimates of reliabilities of $\hat{\pi}$, after a classical bootstrapping with replacement. We tested whether the reliability of the estimate in all cues condition is larger than the best single cue alone (effective cue combination, $\hat{r}_{all} > \hat{r}_{max}$) and whether equal to the value predicted by the optimal cue combination rule ($\hat{r}_{all} = \hat{r}_0$) at $p = .05$ level. The z score (1): Comparison with best individual cue; z score (2): comparison with optimal cue combination. The results indicate that the empirical reliabilities in all cues condition are all larger than the best single cue alone but lower than predicted by optimum combination rule, except for IB in the light from the left condition.

where we number the three cues in the order of cast shadows, shading, and specular highlight for convenience. The minimum variance estimate of $\hat{\pi}$ when all three cues are present is the weighted convex combination of each single cue, which is given by

$$\hat{\pi}_0 = \sum_{i=1}^{3} W_i \hat{\pi}_i \quad , \quad W_i = r_i \Big/ \sum_{j=1}^{3} r_j, \qquad (6)$$

where $r_i = \sigma_i^{-2}$ is defined as the reliability of that cue (terminology due to Backus & Banks, 1999). If the cues are combined according to Equation 6, then the reliability of the combined estimate is the sum of individual reliabilities $r_0 = \sum_{i=1}^{3} r_i$. We can derive similar equations for estimates of ψ_P^j based on multiple cues.

We first want to test whether, in the all cues condition, observers combined information from multiple (two or more) cues to produce an estimate that is more reliable than the estimates based on any one of the cues. That is, we want to test whether

$$r_{all} > \max\{r_1, r_2, r_3\}, \qquad (7)$$

where r_{all} denotes the observer's performance with all cues present. This condition (see Oruç et al., 2003) is the minimum condition that an observer's performance should satisfy if we are to claim that s/he is combining cues effectively and not, for example, using one cue on some trials and

another on the remainder. We refer to Equation 7 as the effective cue combination hypothesis. This condition is strictly weaker than the optimal cue combination hypothesis, which requires $r_{all} = r_0$. We estimated reliabilities for each observer in each cue condition by bootstrap simulation (Efron & Tibshirani, 1993). The resulting estimates are given in Table 2 (for $\hat{\pi}$) and Table 3 (for $\hat{\psi}_P$).

First, we tested the effective cue combination hypothesis for π and ψ_P estimates. We compared $\hat{r}_{max} = \max\{\hat{r}_1, \hat{r}_2, \hat{r}_3\}$, the maximum reliability of any of the single cues conditions, to the reliability $\hat{r}_{all}$ of the all cues present condition (empirical). We computed the z score

$$z = \frac{\hat{r}_{max} - \hat{r}_{all}}{\sqrt{var(\hat{r}_{max}) + var(\hat{r}_{all})}}. \qquad (8)$$

Using the z scores, we tested the hypothesis that $\hat{r}_{max}$ and $\hat{r}_{all}$ come from the same distribution at the .05 chance level. We conclude that both observers used at least two cues because the z scores suggest that the $\hat{r}_{all}$ estimates lie above $\hat{r}_{max}$ and outside the 95% confidence interval for π estimates. The z scores are shown in Tables 2 and 3.

Next, we computed an estimate of the optimal reliability $\hat{r}_0$ as predicted by Equation 6. We compared this estimate to the reliability, $\hat{r}_{all}$, of the all cues condition (empirical) in a manner similar to our previous analysis. We rejected the hypothesis of optimal cue combination for both observers in all conditions except for observer IB in light from the left condition in case of combining cues for a π estimate.

Cue condition	Observers			
	KD—L	KD—R	IB—L	IB—R
All cues	.016	.015	.027	.008
$\hat{r}_{max}$	.010	.008	.009	.009
$\hat{r}_0$—Predicted optimum	.025	.013	.018	.022
Cast shadows	.011	.008	.003	.009
Shading	.010	.004	.005	.007
Highlights	.005	.002	.009	.005
z score (1)	−4.90	−8.22	−10.86	2.83
z score (2)	8.56	−2.28	−5.98	19.18

Table 3. Computed and empirical estimates of reliabilities of $\hat{\psi}$. Same analysis procedure as in Table 2. In all cases, empirical reliabilities in all cues condition are larger than the best individual cue but lower than predicted by optimum cue combination rule at $p = .05$ level, except in three cases: for observer IB under the light from the right, the effective cue combination is violated ($\hat{r}_{all} < \hat{r}_{max}$), and for observer KD under the light from the right and for observer IB under the light from the left, the reliability of all cues condition exceeds the value predicted by the optimum cue combination.

In summary, we found that both observers combined cues effectively but not optimally: $\hat{r}_{max} < \hat{r}_{all} \leq \hat{r}_0$; equality was achieved in only one out of eight conditions.

Summary and discussion

We know human observers take into account the spatial and spectral distribution of light sources in a scene in judging surface albedo and chromaticity (Boyaci et al., 2003, 2004; Ripamonti et al., 2004; Snyder et al., 2005). To do so, a visual system must obtain information about the spatial distribution of light sources in the scene from the scene itself (the lighting model). We tested three possible cues, each of which could provide information about the spatial distribution of light sources in a scene: (1) cast shadows, (2) shading of matte objects, and (3) the virtual images formed in specular objects ("highlights"). We tested observers in scenes with only one cue type available and in scenes with all cues present. In this paper, we measured how accurately observers estimate the parameters of the lighting model from each of the three candidate cues. We do not claim that the parameterization of the lighting model we have chosen is the one used by the visual system, only that whatever parameterization the visual system uses can represent the range of scenes captured by π (punctate-total balance) and Υ_P (light source direction). Our argument and analyses would not be affected by a change of parameterization (Appendix B).

Given the logic of null hypothesis testing, we cannot conclude that an observer made no use of a cue when it was available if we fail to reject the corresponding null hypothesis. We can only note those cases where we found evidence, in the form of a p value, that a particular observer used a particular cue. We concluded that, in the scenes containing only one cue type, at least some observers used each of the cues to estimate the lighting model in Experiment 1. In contrast, we reject this null hypothesis for the shading cue for only one observer. This outcome may simply indicate that the shading cue elements we provided ("boxes") are not very good shading cues.

Furthermore, our results indicate that there are individual differences in cue use. Such individual differences in cue use are commonly found in the cue combination literature (Oruç et al., 2003). Individual differences abound in depth perception studies (for example, the widely ranging cue reliabilities across subjects found by Hillis, Ernst, Banks, & Landy, 2002, or the large variation in abathic distance found by Kontsevich, 1998). Moreover, recent studies show that associative learning can alter perceptual appearance (Adams, Graf, & Ernst, 2004; Jacobs & Fine, 1999; Sinha & Poggio, 2002) leading to the suggestion that cues can be learned (Backus, 2005). If at least some cues are learned, then we would expect individual differences as a result of differences in perceptual experience.

Our findings in this study suggest that the information about spatial organization is useful in estimating the lighting model of the scene and eventually surface albedo. Given those results, it is plausible that the visual system uses the information that it derives from the cues to the spatial organization of the illumination in estimating the lighting model and surface albedo (or chromaticity).

In the second experiment, we tested whether observers combined cues when more than one is present in the scene. We used a criterion that we referred to as effective cue combination: the observer must combine available cues into an estimate that is more reliable than any of the cues in isolation (Oruç et al., 2003). We rejected the null hypothesis of no effective cue combination. Observers were able to combine information from multiple cues to produce estimates of the punctate-total ratio $\hat{\pi}$ and $\hat{\psi}_P$ (the parameters characterizing the lighting model) that were more reliable than any of the three estimates that they derived in the single cue conditions.

We also tested the hypothesis that observers combined cues optimally. We rejected the hypothesis of optimal cue combination for all (total of eight: two observers × two light sources × two parameter estimates) but one condition. This outcome is unsurprising. It is implausible that any biological visual system performs any visual task optimally, just as it is implausible that any coin is precisely "fair." A finding of "optimality" involves failing to reject the hypothesis that the observer's performance is optimal and could be the result of a Type II error. A similar experiment with more experimental trials (statistical power) would likely reverse a finding of optimality. The criterion of effective cue combination, introduced in Oruç et al. (2003), is more useful in examining how biological organisms combine cues and in determining whether they do so.

Appendix A

Combining *p* values

For each cue condition, we tested the null hypothesis that none of the observers had made use of the cue(s) present in the condition. For each observer in each condition, we computed the exact p value for a test of the hypothesis that this observer is using the cue(s) available, reported in Table 1. It can be seen that several of these values are very small. If we tested each of these values against the .05 level, we would simply record whether the exact value was less than or equal to .05 (reject) or greater than .05 (not reject). We wanted to combine this information. We made use of the following: if the null hypothesis was true, then the distribution of each of these exact p values would be uniform on the interval [0,1]. That is, we would expect the exact p value to be greater than .5 one half of the time.

For each cue condition, we computed as a test statistic the mean of the logarithm of the exact p values for each of
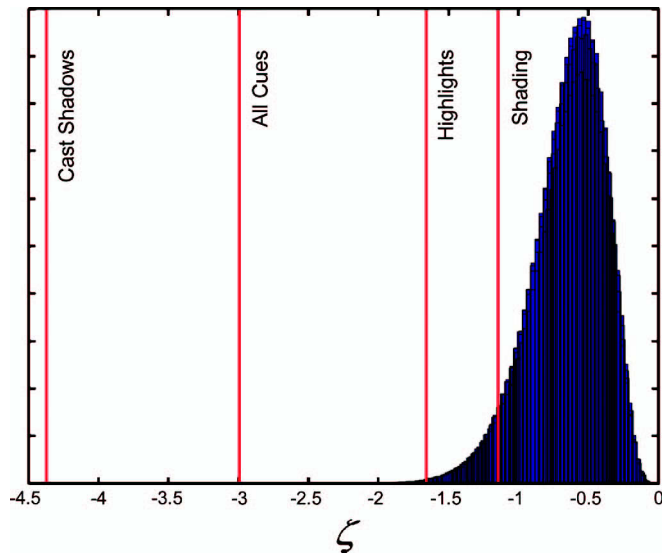
Figure 11. The distribution of the mean of log *p* values. For each cue condition, we plot a vertical red line at $\log_{10}\bar{p}$, the logarithm of the geometric mean across subjects of the *p* values in Table 1. Under the null hypothesis that no subject is making use of the available cue(s), the distribution of $\log_{10}\bar{p}$ follows the blue histogram shown. For three of the four conditions, $\log_{10}\bar{p}$ is far outside its expected range and we may reject the null hypothesis. For the remaining cue condition (shading), the value of the statistic is barely significant at the .05 level.

the six observers, $p_k = 1, 2,\ldots, 6$,

$$\zeta = \sum_{k=1}^{6} \log_{10}(p_k)/6. \tag{A1}$$

Figure 11 contains a histogram of the distribution of ζ under the null hypothesis that none of the observers is making use of the cue. If this hypothesis was true for any cue condition, we would expect the computed value ζ to fall above the 5th percentile of the histogram with probability .95. If, however, the *p* values for the test tend to be small for some of the observers, then the test statistic ζ will tend to be too small, falling in the extreme left tail of the histogram. The value of the test statistic ζ for each cue condition is marked with vertical red lines. We reject the null hypothesis for all cue conditions at the .05 level. We can report the exact *p* value of the test statistic ζ for each cue condition, just as we can for any test statistic whose distribution under the null hypothesis is known. These values are reported in the rightmost column of Table 1.

Appendix B

Reparameterization

The parameter estimate $\hat{\pi}$ is only defined on the interval [0,1] and its distribution near 0 and near 1 is highly skewed.

Ordinary confidence intervals, based on Gaussian distributions, (e.g., ±2 *SD*) could include values outside the range [0,1] that are not physically possible. To avoid such difficulties we performed the analysis with the following change of variable

$$\gamma = \log\left[\frac{\pi}{1-\pi}\right], \tag{B1}$$

whose range includes the entire real line $(-\infty, \infty)$. The function in Equation B1 has inverse

$$\pi = \frac{e^{\gamma}}{1 + e^{\gamma}}. \tag{B2}$$

Using Equations B1 and B2, we can translate back and forth from one parameterization to the other.

All maximum likelihood estimates of confidence interval limits and all hypothesis tests were computed in terms of the γ parameterization and then transformed into the π parameterization by Equation B2. The justification for using these procedures is that maximum likelihood estimation and hypothesis testing based on likelihood ratios are invariant under invertible re-parameterizations such as the ones above (Pawitan, 2001, pp. 43–45). For example, if $\hat{\pi}$ is the MLE of π in one parameterization, then its transformation $\hat{\gamma} = \log[\hat{\pi}/1-\hat{\pi}]$ will be equal to the MLE $\hat{\gamma}$ if we instead maximized likelihood in the other.

Acknowledgment

Commercial relationships: none.
Corresponding author: Huseyin Boyaci.
Email: boyac003@umn.edu.
Address: Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA.

Footnotes

[1]So that there would be no visible variation in shading across the surface of the poles themselves.
[2]In designing the stimuli, we effectively superimposed the virtual space containing the stimuli and the actual space containing the observer. We speak of distances and locations in the two spaces interchangeably.

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7,* 1057–1058. [PubMed]

Backus, B. (2005). *Recruitment of new cues for visual perception through Pavlovian conditioning*. Presentation at the European Conference on Visual Perception, A Coruna, Spain, August 22–26.

Backus, B. T., & Banks, M. S. (1999). Estimator reliability and distance scaling in stereoscopic slant perception. *Perception, 28,* 217–242. [PubMed]

Bloj, M., Ripamonti, C., Mitha, K., Hauck, R., Greenwald, S., & Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision, 4*(9), 735–746, http://www.journalofvision.org/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), 664–679, http://www.journalofvision.org/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), 541–553, http://www.journalofvision.org/3/8/2/, doi:10.1167/3.8.2. [PubMed] [Article]

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman-Hall.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433. [PubMed]

Gilchrist, A. L. (1977). Perceived lightness depends on spatial arrangement. *Science, 195,* 185–187. [PubMed]

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

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 106,* 795–834. [PubMed]

Hara, K., Nishino, K., & Ikeuchi, K. (2005). Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27,* 493–505. [PubMed]

Haralick, R. M., & Shapiro, L. G. (1993). *Computer and robot vision, Vol. 2* (chap. 12). Reading, MA: Addison-Wesley.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298,* 1627–1630. [PubMed]

Jacobs, R. A., & Fine, I. (1999). Experience-dependent integration of texture and motion cues to depth. *Vision Research, 39,* 4062–4075. [PubMed]

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shad(ow)ed random Gaussian surfaces. *Perception, 33,* 1405–1420. [PubMed]

Kontsevich, L. L. (1998). Defaults in stereoscopic and kinetic depth perception. *Proceedings of the Royal Society of London. Series B, 265,* 1615–1621. [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. [PubMed] [Article]

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412. [PubMed]

Larson, G. W., & Shakespeare, R. (1996). *Rendering with radiance: The art and science of lighting and visualization*. San Francisco: Morgan Kaufmann.

Li, Y., Lin, S., Lu, H., & Shum, H. (2003). Multiple-cue illumination estimation in textured scenes. *Proceedings of the IEEE International Conference on Computer Vision,* 1366–1373.

Maloney, L. T. (2002). Illuminant estimation as cue combination. *Journal of Vision, 2*(6), 493–504, http://www.journalofvision.org/2/6/6/, doi:10.1167/2.6.6. [PubMed] [Article]

Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics, 3rd Edition*. New York: McGraw-Hill.

Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research, 43,* 2451–2468. [PubMed]

Pawitan, Y. (2001). *In all likelihood: Statistical modeling and inference using likelihood*. New York: Oxford.

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

Scheifler, R. W., & Gettys, J. (1996). *X Window system: Core library and standards*. Boston: Digital Press.

Sinha, P., & Poggio, T. (2002). High-level learning of early perceptual tasks. In M. Fahle (Ed.), *Perceptual learning*. Cambridge, MA: MIT Press.

Snyder, J. L., Doerschner, K., & Maloney, L. T. (2005). Illumination estimation in three-dimensional scenes with and without specular cues. *Journal of Vision, 5*(10), 863–877, http://journalofvision.org/5/10/8/, doi:10.1167/5.10.8. [PubMed] [Article]

von Helmholtz, H. (1866/1924). *Helmholtz's treatise on physiological optics, Volume 1*. In J. P. C. Southall (Ed. & Trans.). New York: The Optical Society of America (Translated from the 3rd German edition).

Wang, Y., & Samaras, D. (2003). Estimation of multiple directional light sources synthesis of augmented reality images. *Graphical Models, 65,* 185–205. [Article]

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