# Lightness constancy and illumination discounting

Alexander D. Logvinenko · Rumi Tokunaga

Published online: 18 June 2011 © Psychonomic Society, Inc. 2011

**Abstract** Contrary to the implication of the term "lightness constancy", asymmetric lightness matching has never been found to be perfect unless the scene is highly articulated (i.e., contains a number of different reflectances). Also, lightness constancy has been found to vary for different observers, and an effect of instruction (lightness vs. brightness) has been reported. The elusiveness of lightness constancy presents a great challenge to visual science; we revisit these issues in the following experiment, which involved 44 observers in total. The stimuli consisted of a large sheet of black paper with a rectangular spotlight projected onto the lower half and 40 squares of various shades of grey printed on the upper half. The luminance ratio at the edge of the spotlight was 25, while that of the squares varied from 2 to 16. Three different instructions were given to observers: They were asked to find a square in the upper half that (i) looked as if it was made of the same paper as that on which the spotlight fell (lightness match), (ii) had the same luminance contrast as the spotlight edge (contrast match), or (iii) had the same brightness as the spotlight (brightness match). Observers made 10 matches of each of the three types. Great interindividual variability was found for all three types of matches. In

particular, the individual Brunswik ratios were found to vary over a broad range (from .47 to .85). That is, lightness matches were found to be far from veridical. Contrast matches were also found to be inaccurate, being on average, underestimated by a factor of 3.4. Articulation was found to essentially affect not only lightness, but contrast and brightness matches as well. No difference was found between the lightness and luminance contrast matches. While the brightness matches significantly differed from the other matches, the difference was small. Furthermore, the brightness matches were found to be subject to the same interindividual variability and the same effect of articulation. This leads to the conclusion that inexperienced observers are unable to estimate both the brightness and the luminance contrast of the light reflected from real objects lit by real lights. None of our observers perceived illumination edges purely as illumination edges: A partial Gelb effect ("partial illumination discounting") always took place. The lightness inconstancy in our experiment resulted from this partial illumination discounting. We propose an account of our results based on the two-dimensionality of achromatic colour. We argue that large interindividual variations and the effect of articulation are caused by the large ambiguity of luminance ratios in the stimulus displays used in laboratory conditions.

A. D. Logvinenko (⋈) Department of Vision Sciences, Glasgow Caledonian University, City Campus, Cowcaddens Road, Glasgow G4 0BA, UK e-mail: a.logvinenko@gcu.ac.uk

R. Tokunaga Research Institute of Electrical Communication, Tohoku University, Katahira Aoba, Sendai 980-8577, Japan

Springer

Keywords Color and light · Lightness/Brightness · Surface brightness · Individual differences · Illumination edge · Luminance contrast · Intrinsic images · Gelb effect · Articulation · Brunswik ratio

A remarkable feature of human vision, which still remains to be understood, is lightness constancy with respect to illumination (for reviews, see, e.g., Gilchrist, 2006; Katz, 1935). Such constancy poses a problem if one believes, as many visual scientists do, that lightness is derived somehow from reflected light that is the only input into the human visual system. The problem is that reflected light intensity is proportional to both surface albedo and the intensity of the light falling onto this surface. How, then, does the visual system derive (from the reflected light) lightness that correlates with surface albedo, despite changes in incident illumination?

A textbook explanation suggests that the visual system simply discounts the illumination changes due to adaptive mechanisms of gain control in photoreceptors and other visual neurons (e.g., Cornsweet, 1970; Whittle, 1994). Responding to luminance contrast rather than to luminance itself, such neurons can indeed encode reflectance (i.e., albedo) irrespective of the ambient light intensity (e.g., Shapley & Enroth-Cugell, 1984). However, there are a few problems with this explanation. First, it predicts complete lightness constancy. However, one of the most puzzling problems concerning lightness constancy is that, when measured, it is almost never perfect<sup>1</sup> (Gilchrist, 2003; Katz, 1935; Woodworth, 1938, pp. 493–499). The failures of lightness constancy as measured in many laboratory experiments were so considerable and systematic that they led Gilchrist (2003) to claim that these "errors" in lightness perception presented even more fundamental a problem than veridical lightness perception.

Second, this explanation works only for homogeneous illumination, because it implicitly interprets all luminance edges in the scene as lightness edges (i.e., those produced by material edges). That is, the discounting-illumination hypothesis implies that lightness is the only output in response to reflected light entering the eye. Being applied to a scene containing illumination and material borders, such a mechanism would mistake all of the illumination borders for material ones. In other words, the discounting-illumination hypothesis predicts discounting of illumination edges at the phenomenological level (i.e., perceiving them as lightness edges) under inhomogeneous illumination. Erroneous perception of this kind may happen in human vision. A classical example is the so-called Gelb effect (Gelb, 1929), in which the complete discounting of illumination takes place. Another example is misperceiving a shadowed area as an ink stain (e.g., Logvinenko & Menshikova, 1994). However, these illusory perceptions are rather rare in everyday life.

So, if our perception of natural scenes operated as the discounting-illumination hypothesis predicts, we would not be able to perceive illumination changes in space or/and time. However, we are capable of perceiving both illumination and lightness. Shadows and other spatial inhomogeneities of illumination are as important features of any complex scene in

the natural environment as material objects. It is vitally important from the ecological point of view to distinguish between these (e.g., Gibson, 1979). Human observers have been found to be rather good at distinguishing between changes produced by illumination and reflectance in visual scenes (e.g., Craven & Foster, 1992; Foster et al., 2001; Foster, 2003; Kingdom, 2008).

Although it seems self-evident that illumination must be represented in our perception, there is no consensus concerning the phenomenological status of illuminationthat is, the perceptual correlate of illumination. The most common view is that there are two perceptual dimensions, apparent illumination and lightness, the former correlating with illumination and the latter with reflectance (Gilchrist, 2006). The problem with this view is that it is very hard (especially for inexperienced observers) to assess the apparent illumination of a particular object. At best, they can judge the level of ambient illumination rather than local apparent illumination. Much better is our capacity to compare the levels of illumination of two differently illuminated areas (e.g., highlight vs. shadow). However, subjective estimates of apparent illumination are rather rough and subject to large uncertainty. This is in dramatic contrast with lightness, which is distinctively perceived for every fragment of a surface and can be subjectively estimated with much higher certainty. Notably, there is no commonly accepted scale for measuring apparent illumination comparable with, for instance, the Munsell lightness scale.

All this has led to the view that apparent illumination does not exist as a perceptual dimension per se. Instead, we are aware of the level of illumination as being due to inferring rather than perceiving. For instance, MacLeod (1940) believed that apparent illumination is inferred from the brightness (i.e., perceived luminance) of the light coming to the eye. However, this seemingly simple idea leaves us with a paradox. Apparent illumination is supposed to be directly inferred from brightness (i.e., perceived luminance). Lightness is supposed to be computed indirectly from luminance. However, lightness is more easily and accurately perceived than apparent illumination. Yet, if not brightness, what is apparent illumination inferred from?

We argue that apparent illumination might be inferred from lightness. Recall that *lightness* usually refers to a perceptual continuum of achromatic colours that vary from white through various shades of grey to black. However, as has been noticed by a number of previous authors (e.g., Katz, 1935), achromatic colours constitute a one-dimensional continuum only when illumination is fixed. In a real scene with multiple light sources, achromatic colours vary along two different perceptual dimensions. Katz called them *lightness* and *pronouncedness*; Koffka (1935), *whiteness* and *brightness*. The former correlates with reflectance, the latter with illumination. That the human visual system produces two-dimensional perceptual

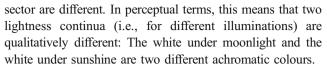


<sup>&</sup>lt;sup>1</sup> Notably, complete lightness constancy has been found only in studies in which well-articulated stimulus displays were employed (e.g., Burzlaff, 1931). In the present context, *articulation* means the number of surfaces with different albedos (Gilchrist & Vidal, 2002).

output in response to the one-dimensional input (reflected light intensity) has been supported by a study that confirmed that achromatic colours composed a two-dimensional manifold (Logvinenko & Maloney, 2006). The achromatic colour dimensions that emerged in this study from multidimensional analysis were termed *lightness* and *surface brightness*.<sup>2</sup> It was shown that lightness was mainly related to surface albedo, and surface brightness to surface illumination. It should be emphasised that surface brightness was found to covary with incident rather than reflected light. There was ample evidence from this and related studies (Logvinenko, 2005a, 2005b) that surface brightness is different from the brightness of the reflected light.

One might ask: Is surface brightness just a euphemism for apparent illumination? The answer is negative, if lightness and apparent illumination are thought of as one-dimensional perceptual correlates of their physical counterparts (i.e., albedo and light intensity, respectively). In this case, the achromatic colour of an object could be specified by the lightness and apparent illumination values as a rectangle is by its width and height. Speaking more formally, in this case the achromatic colour could be considered as a Cartesian product of the lightness and apparent-illumination dimensions. However, it has proven that the manifold of achromatic colours cannot be represented as a Cartesian product of two onedimensional continua, those of lightness and surface brightness.<sup>3</sup> In other words, the achromatic colour manifold *cannot* be specified by lightness and surface brightness as a rectangle is by its width and height. More specifically, it was found that the lightness continuum shrank when the light intensity decreased (Logvinenko & Maloney, 2006). Particularly, the subjective distance between black and white decreases under dimmer light. Therefore, a more appropriate geometrical representation for the achromatic colour manifold is a circular sector, where the radii represent surface brightness, and the arcs, lightness (Logvinenko & Maloney, 2006). Thus, complete darkness is represented by the circle centre, where the lightness continuum collapses into a point. A circular sector is obviously not a Cartesian product of a radius and an arc. Indeed, all of the arcs constituting the

The existence of surface brightness was also revealed for chromatic Munsell papers using the same multidimensional scaling technique (Tokunaga, Logvinenko, & Maloney, 2008). Furthermore, it turns out that surface brightness is just one of the three additional (so-called *lighting*) dimensions of object colour that have been recently described along with the traditional (so-called *material*) dimensions of object colour (Tokunaga & Logvinenko, 2010a, 2010b, 2010c).



A remarkable fact concerning surface brightness is that, on the one hand, it is easy to see the phenomenological difference between two areas of different illuminations in a natural scene. The notion of surface brightness appeals to this difference. On the other hand, when the scene is homogeneously illuminated, it is very difficult to say what the surface brightness of a particular object is. In other words, one needs at least one illumination border for surface brightness to emerge.4 Therefore, one can readily judge which of two areas of different illuminations is brighter (or dimmer), but it is hard, if not impossible, to estimate an absolute level of apparent illumination in a single-light scene. In terms of the example with a circular sector, one can say that when only one arc is selected, it is impossible to say what its position is in the sector in the absence of the whole sector. Still, given two arcs, it is easy to make a comparative judgment concerning the arcs' lengths.

Following on from this analogy, one can say that in the same way that a sector can be considered as an ordered family of arcs, the two-dimensional manifold of achromatic colours can be thought as an ordered family of the lightness continua. From this point of view, there is only lightness, and surface brightness is a euphemism of the order on the family of lightness continua.<sup>5</sup> Therefore, in contrast to classical dualistic frameworks that incorporate lightness and apparent illumination as independent perceptual dimensions, we argue in favour of the monistic view (for an elaboration of the controversy between "monism vs. dualism" in lightness perception, see Gilchrist & Economou, 2003): There is nothing but lightness. Yet, one can infer the level of apparent illumination depending on the lightness continuum to which the particular lightness value belongs. However, this is tacit knowledge. We do not believe that apparent illumination is phenomenologically present as directly as lightness.

Thus, the manifold of achromatic colours can be considered as an ordered bundle of one-dimensional continua, each of which is, in turn, subject to another order. The order within a single continuum is referred to as the *lightness order*, the continuum itself being referred to as the *lightness continuum*. The order imposed on the lightness continua as whole is referred to as the *surface-brightness order*.



<sup>&</sup>lt;sup>3</sup> If the manifold of achromatic colours were a Cartesian product of the lightness and surface-brightness continua, then for encoding achromatic colours—say, for the purpose of computer graphics—one would need two look-up-table vectors (one for lightness and one for surface brightness). In the context of this computer-graphic metaphor, the two-dimensionality of achromatic colours means that one needs a look-up-table matrix.

<sup>&</sup>lt;sup>4</sup> Recall that to see lightness, one needs at least one reflectance border as well. Otherwise, one will see the Ganzfeld (i.e., an objectless visual impression).

<sup>&</sup>lt;sup>5</sup> Admittedly, if this order satisfies the axioms of weak order, there exists an ordinal scale for it (Krantz et al., 1971) that can be interpreted as a surface-brightness scale. Still, the existence of such a scale does not mean that surface brightness is embodied as a perceptual dimension on its own.

In a natural scene, each luminance border is perceived as a border dividing the areas of different achromatic colours, which can be different with respect to both lightness and surface brightness, the surface-brightness difference being understood as containing different lightness continua. It turns out that surface brightness is related to luminance in a different way than lightness is. Specifically, it has been noticed that when a luminance edge produced by illumination is misperceived as a lightness edge, a sort of contrast enhancement takes place (Hering, 1964). For example, when a shadow appears to be a black (lightness) spot due to an apparent depth reversal, it looks much more pronounced (Logvinenko & Menshikova, 1994). Also, it has been found that luminance contrast produced by a sharp illumination edge is underestimated by a factor of 2.2, as compared to the same luminance contrast produced by a reflectance edge (Logvinenko, 2005b). Even larger underestimation of illumination was revealed by Logvinenko and Maloney (2006). They discovered that the dissimilarity between two different grey papers equally illuminated was judged to be 16 times as much as that between two identical grey papers illuminated by two different lights, despite the fact that the luminance difference between the two papers was the same in both cases (Logvinenko & Maloney, 2006). These results can be understood if one assumes that the lightness-versus-reflectedlight-intensity response function is steeper than the surfacebrightness-versus-reflected-light-intensity response function. The fact that the same luminance contrast produces more contrast in terms of lightness than of surface brightness will be referred to as illumination contrast discounting.

It should be mentioned that illumination contrast discounting has been found to be subject to large interindividual variability (Logvinenko, 2005b). As pointed out above, experimental measurements of lightness constancy have also exhibited interindividual differences (e.g., Katz, 1935; Thouless, 1932; Woodworth, 1938). "All authors who have performed experiments on colour constancy have been able to confirm the fact that individual differences are extraordinarily high in comparison with those in other visual experiments ..." (Katz, 1935, p. 137). It must be said, however, that although the discord in lightness judgements observed in these studies was quite impressive, only few observers participated (e.g., 4 in Katz's experiment, or 7 in Thouless's). If confirmed, the individual "errors" in lightness matching could be an even greater challenge for any lightness perception theory than the imperfectness of lightness constancy obtained in the many previous experiments.

Consequently, we have deliberately employed a much more representative group of observers in the following experiment, which was designed (1) to investigate the individual variations in lightness constancy and illumination contrast discounting and (2) to ascertain whether these two phenomena are quantitatively related. Specifically, we have measured both lightness constancy and

illumination contrast discounting for the same observers and in the same experimental conditions, and correlated their responses.

## Method

In all, 44 observers (including both authors) participated in the experiments. As vision science department students and staff, all observers were familiar with the notions of luminance and luminance contrast, but except for the authors, they were unaware of the purpose of the experiment. All had normal or corrected-to-normal vision.

The observers sat in a chair that was 1.3 m from the stimulus display (55 × 60 cm). They were presented with a sheet of black (background) paper mounted on a black wall<sup>6</sup> dimly illuminated (1.5 cd/m<sup>2</sup>) by a projector (L1 in Fig. 1). Another projector (L2 in Fig. 1) was also used to produce a rectangular spotlight (37.8 cd/m<sup>2</sup>). Thus, there was a horizontal illumination edge across most of the background (Fig. 1). The luminance ratio made at this illumination edge was 25. Two arrays, each of 20 squares (3.1 cm each) of different albedos (thus, greyness), were printed on the shadowed half of the background ("target squares" in Fig. 1). The arrays in the left and right panels were identical (in particular, they had identical albedos), except for being randomised in location. The albedo of the target squares varied in approximately equal steps, so that the luminance ratio (between them and the background) varied from 2 to 16.1.7

Several measures were taken to make the experimental situation seem as close to natural as possible. Vision was binocular and no time restriction was imposed. The experimental room was semidarkened. While the edges of the spotlight were sharp, there were plenty of illumination cues pointing out that the horizontal luminance border was produced by illumination. Nevertheless, a strong Gelb effect—that is, seeing a spotlight as a pigmented area—was reported by some observers. Thus, a series of 12 grey squares (with luminance ratios from 2.3 to 16.9) were introduced at the bottom of the spotlight, to increase articulation in order to reduce the Gelb effect ("articulating squares" in Fig. 1).

Luminance measurements were made with a Minolta photometer (Model LS-110). To evaluate the luminance contrast of the target squares, five luminance measurements were made for each square—specifically, the luminance of the light reflected from the central area of the square itself,  $L_{\rm Sq}$ , and that of its immediate surround at the top, bottom,

<sup>&</sup>lt;sup>7</sup> This was the maximal luminance ratio that we could achieve using the printing facilities available.



<sup>&</sup>lt;sup>6</sup> The walls in the experimental room were dark grey. The part of the wall in front of the observer was covered with black cloth.

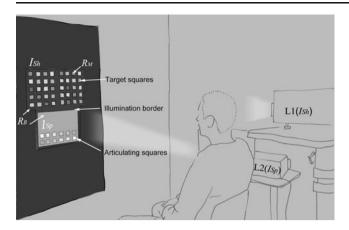


Fig. 1 Experimental display (see text)

right, and left sides of the square. The average of the last four measurements was taken as the luminance of the surround,  $L_{\rm Sr}$ . The luminance contrast of the square was specified as  $(L_{\rm Sq}-L_{\rm Sr})/L_{\rm Sr}$ 

Three experiments were conducted (the main experiment and two controls). All of the 44 observers took part in the main experiment, which was divided into two parts. In the first part, the task was to make a *contrast match*. Observers were asked to find which of the target squares had the same contrast as (or the closest contrast to) that of the horizontal illumination edge.

More specifically, in the very beginning of the experiment, observers were asked, "You see this spotlight here, do you not?" Nobody expressed the slightest doubts about the presence of the spotlight. Then the experimenter asked whether they could find a target square that had the same luminance contrast as the one that the illumination border made with the background. None said "no." All of them chose some target square.

In the second part of the experiment, the observers had to make a *lightness match*. The observers were asked, first, "Do you see the background in both the highlighted and the shadowed regions as made from the same paper?" All of them answered negative. The negative answer testifies that the observers based their response not on what they knew to be true about the display but rather on its perceptual image. This justified the following instruction they were then suggested: to choose a target square in the shadowed area that *looked as if* it was made of the same paper as the upper part of the highlighted region.<sup>8</sup>

To ascertain the effect of the articulating squares, we conducted the first control experiment, which was identical to the main experiment except that the spotlight was void of

<sup>&</sup>lt;sup>8</sup> Such an instruction had been used in many previous studies on lightness constancy (Gilchrist, 2006, p. 7). It prevents experimenters from needing to explain to naive observers what lightness is.



the articulating squares. Only 34 of the 44 observers who participated in the main experiment took part in the first control experiment. This was because the main experiment lasted a few months, and some of the observers were unavailable for the control observations.

Also, the second control experiment was conducted with 23 observers who had participated in the main and the first control experiments. In the second control experiment, observers were instructed to *match the brightness* of the spotlight. The rationale was to find out the extent to which the observers' matches depended on the instructions. The second control experiment was carried out for both high articulation (i.e., with articulating squares in the spotlight) and low articulation (i.e., without articulating squares).

In all three of the experiments, in one trial 2 matches (from each of the two arrays of target squares) were made. Then the experimenter replaced the whole sheet with the other one. The squares in the upper (shadowed) part were arranged in a random order on each sheet. In all, six sheets were used for each task. The matches made with the first sheet were then discarded. Therefore, 10 matches for each task (i.e., contrast, lightness, and brightness matching) and for each condition (i.e., low and high articulation) were scored for each observer in each experiment.

To sum up, 44 observers conducted 20 matches each: 10 contrast matches and 10 lightness matches under high-articulation conditions (in the main experiment). Of these, 34 observers made another 20 matches, thus in total 40 matches each: 10 contrast and 10 lightness matches under both high- (the main experiment) and low- (the first control experiment) articulation conditions. Then, 23 of the observers made 60 matches each: 10 contrast, 10 lightness, and 10 brightness matches under both high- and low-articulation conditions. In total, 2,020 matches were recorded.

## Results

The distributions of all of the contrast and lightness matches obtained in the main experiment are presented in Fig. 2. The contrast matches fall into the luminance contrast range from 1.7 to 15.1, with the mean being 8.4. Therefore, for the luminance contrast at the material border to match the luminance contrast at the illumination border, the latter should exceed the former by some factor that we will refer to as the *illumination contrast discounting index*. The luminance contrasts of the match square and the spotlight can be expressed as  $C_{\rm M} = (I_{\rm Sh}R_{\rm M} - I_{\rm Sh}R_{\rm B})/I_{\rm Sh}R_{\rm B}$  and  $C_{\rm I} = (I_{\rm Sp}R_{\rm B} - I_{\rm Sh}R_{\rm B})/I_{\rm Sh}R_{\rm B}$ , respectively, where  $I_{\rm Sh}$  and  $I_{\rm Sp}$  are the light intensities of the shadowed background and the spotlight, respectively;  $R_{\rm B}$  and  $R_{\rm M}$  are the albedos of the

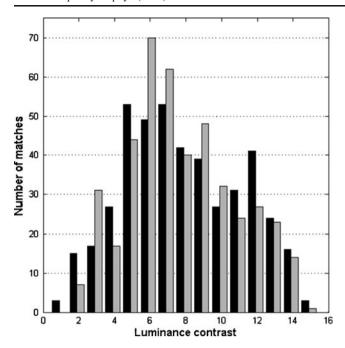


Fig. 2 Histograms of contrast (black) and lightness (grey) matches

background and the matched square. The illumination contrast discounting index,  $C_1/C_M$ , is given by the ratio

$$a = (I_{Sp}R_B - I_{Sh}R_B)/(I_{Sh}R_M - I_{Sh}R_B). \tag{1.1}$$

The mean illumination contrast discounting indices for 44 observers are presented in Fig. 3. They vary from 1.9 to 10.1, with a mean of 3.4. Therefore, on average, compared to the illumination edge, reflectance edges were 3.4 times as effective at producing apparent luminance contrast.

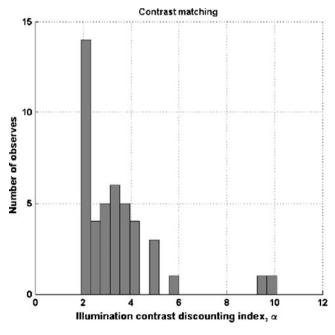


Fig. 3 Histogram of illumination contrast discounting indices (see text)

The luminance contrast range for lightness matching was 2.6 to 15.1, with a mean of 8.2. The boxplot in Fig. 4 compares the lightness matches in samples from all of the 44 observers. As one can see, the individual luminance contrast ranges for lightness matching are considerably smaller than those for the whole group. This shows that the individual differences between observers are real.

A Friedman test showed no significant difference between contrast and lightness matching ( $\chi^2 = 0.52$ , p = .47). Still, for 12 of 44 observers, the lightness match significantly differed from the contrast match according to the Kruskal–Wallis test (p < .05). Yet, even for these 12 observers, there was a strong correlation between the two types of matches.

Figure 5 presents a plot of the mean lightness matches against the mean contrast matches for 44 observers and a linear regression of the former on the latter. A quadratic regression was also performed by modelling the relationship between lightness and contrast matches by a parabola:

$$C_{LM} = k_0 + k_1 C_{CM} + k_2 (C_{CM})^2 + \varepsilon,$$
 (1.2)

where  $C_{\rm LM}$  and  $C_{\rm CM}$  stand for luminance contrasts of lightness and contrast matches, respectively, and  $\varepsilon$  is a random error. A least-squares estimation of the regression coefficients yielded  $\widetilde{k}_0 = 1.99$ ,  $\widetilde{k}_1 = 0.72$ , and  $\widetilde{k}_2 = 0.0026$ . We tested a statistical hypothesis that  $C_{\rm LM} = C_{\rm CM}$ —that is,

$$k_0 = 0, k_1 = 1, k_2 = 0.$$
 (1.3)

The following statistic was evaluated:

$$F = \frac{(S_0 - S)/3}{S/(n-3)},\tag{1.4}$$

where S is the quadratic regression sum of the squares:

$$S = \sum\nolimits_{i=1}^{n} \left\{ C_{\mathrm{LM}}(i) - \widetilde{k}_{0} - \widetilde{k}_{1} C_{\mathrm{CM}}(i) - \widetilde{k}_{2} [C_{\mathrm{CM}}(i)]^{2} \right\}^{2}.$$

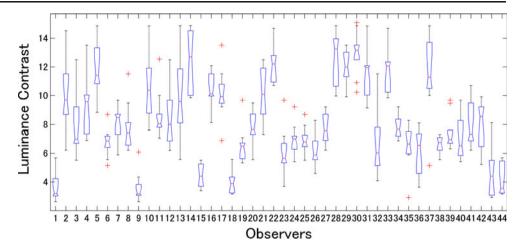
Here,  $C_{\rm LM}(i)$  and  $C_{\rm CM}(i)$  are the mean luminance contrasts set up by the *i*th observer as the lightness and contrast matches, respectively, and n is the number of observers.  $S_0$  is the sum of squares of residuals under the hypothesis in Equation 1.3:

$$S_0 = \sum_{i=1}^n \left[ C_{LM}(i) - C_{CM}(i) \right]^2.$$

We assume that the errors are independent and normally distributed with an expected value of zero and that they all have the same variance. In this case, under the hypothesis (Eq. 1.3), the statistic in Equation 1.4 asymptotically follows an F distribution with 3 and n-3 degrees of freedom. The evaluated value of the statistic (F=2.6) was too small to reject the hypothesis (p=.07). Therefore, the present data do



Fig. 4 Lightness matches for all observers. The vertical axis is the luminance contrast for lightness matching. The line in the middle of each box is the median. The top and bottom of each box mark the 25th and 75th percentiles. The whiskers and the plus signs indicate outliers



not provide statistically significant support for the claim that lightness matches differ from contrast matches.

It should be said that in most previous experiments on lightness constancy, the results were expressed in terms of the so-called *Brunswik ratio* (rather than the luminance contrast of the lightness match):

$$\beta = \frac{I_{Sp}R_B - I_{Sh}R_M}{I_{Sn}R_B - I_{Sh}R_B},\tag{1.5}$$

where  $I_{\rm Sh}, I_{\rm Sp}, R_{\rm B}$ , and  $R_{\rm M}$  are defined as above. In the case of an exact match—that is, when  $R_{\rm M}=R_{\rm B}$ —the Brunswik ratio,  $\beta$ , equals one. When the match is based on the equality of the intensities of the light reflected from the spotlight and the matched square—that is, when  $I_{\rm Sp}R_{\rm B}=I_{\rm Sh}R_{\rm M}$ —the Brunswik ratio equals zero. Figure 6 presents

the histogram of the mean Brunswik ratios obtained in the main experiment for 44 observers. They vary from .47 to .85, with a mean of .66.

Note that Brunswik ratio  $\beta$  (Eq. 1.5) and the illumination contrast discounting index  $\alpha$  (Eq. 1.1) are related as follows:

$$1/a + \beta = 1. \tag{1.6}$$

The reciprocal of the illumination contrast discounting index

$$\frac{1}{\alpha} = \frac{I_{Sh}R_M - I_{Sh}R_B}{I_{Sp}R_B - I_{Sh}R_B},\tag{1.7}$$

shows the extent to which the luminance difference produced by the difference in illumination is less effective

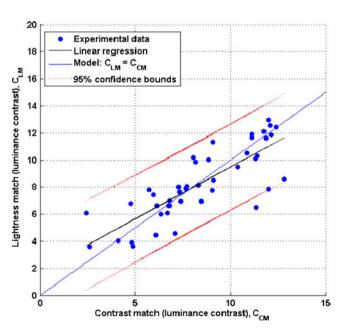


Fig. 5 Linear regression of mean lightness match on mean contrast match

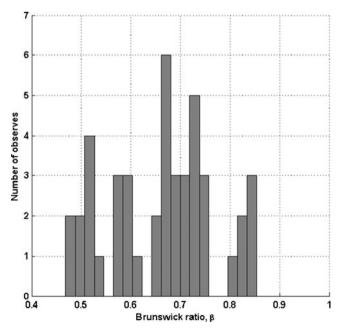


Fig. 6 Histogram of Brunswik ratios (see text)



than that produced by the difference in reflectance. (Recall that  $R_{\rm M}$  in Eq. 1.7 is the albedo of the square that is the contrast match.) Because the illumination contrast discounting index  $\alpha$  varies from 1 to infinity, its reciprocal has the same range as the Brunswik ratio. Figure 7 shows how the Brunswik ratio (Eq. 1.5) and the reciprocal of the illumination contrast discounting index (Eq. 1.7) covary for our 44 observers.

In the first control experiment (i.e., with the low-articulated spotlight), the mean contrast and lightness matches for the 34 observers were 12.9 and 13.1, respectively (vs. means of 8.4 and 8.3 obtained for these observers in the main experiment). A significant difference between contrast and lightness instructions was found for only 1 observer. Figure 8 presents a plot of the reciprocals of the illumination contrast discounting index (and the Brunswik ratios) obtained in the first control experiment (low-articulated display) versus those obtained in the main experiment (high-articulated display). As one can see, all of the points lie above (or, respectively, under) the diagonal of the unit square, indicating an increase of the reciprocal of the illumination contrast discounting index (and, respectively, a decrease of the Brunswik ratio) observed in the first control experiment. It is thus evident that both the apparent luminance contrast of the illumination edge and the lightness of the highlighted background are heightened when the highlighted field is devoid of articulation. Unfortunately, the results exhibit a sheer ceiling effect: The data in Fig. 8 are compressed into a horizontal stretch at the level of approximately. 6 (respectively, .4 for the Brunswik ratio). Since the maximum luminance ratio for the squares (16.1) was smaller than the luminance ratio at the horizontal illumination edge (25), the minimal achievable

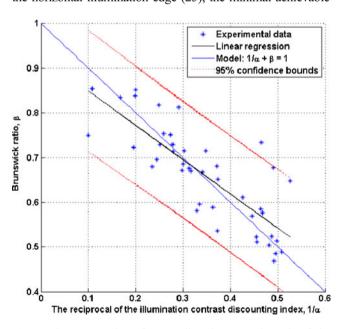


Fig. 7 Linear regression of Brunswik ratios on reciprocals of the illumination contrast discounting index

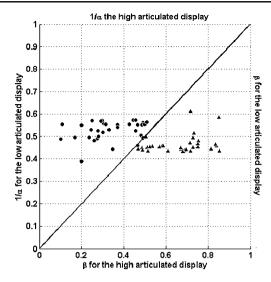


Fig. 8 Effect of articulation on Brunswik ratios (triangles) and on reciprocals of the illumination contrast discounting index (circles)

value of the illumination contrast discounting index was 1.61. Hence, the reciprocal of the illumination contrast discounting index,  $1/\alpha$ , could not have exceeded .62. Likewise, the minimal achievable value of the Brunswik ratio was .38. We believe that this limitation caused the ceiling effect. Therefore, the only conclusion that can safely be drawn from the first control experiment is that both the contrast and lightness matches considerably increased as compared to those obtained in the main experiment.

Brightness matches made by the 23 observers in the second control experiment (Fig. 9) were found to be significantly different, according to a Friedman test, from

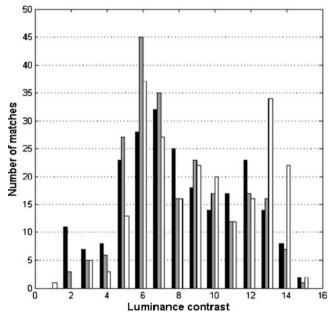


Fig. 9 Histograms of brightness (white), contrast (grey), and lightness (black) matches



both the lightness ( $\chi^2=27.67,\ p<.001$ ), and contrast matches ( $\chi^2=24.53,\ p<.001$ ). Specifically, the mean brightness match was 9.79 (with a range from 1.7 to 15.1), whereas the mean lightness and contrast matches for these 23 observers were 8.56 (with a range from 2.9 to 15.1) and 8.59 (with a range from 2.1 to 15.1), respectively.

Note that while the mean brightness match (9.79) is larger than both the mean lightness and contrast matches, it is considerably smaller than the value predicted from the exact luminance match. The discounting index (Eq. 1.1) evaluated for brightness matches is 3.24. This means that the brightness of the spotlight was underestimated by this factor.

Interestingly, the brightness matches were found to correlate with the lightness matches, actually, the same way as the contrast matches (Fig. 10). Moreover, the brightness matches made for the low-articulated spotlight (i.e., without squares in it) were found to be subject to the same shift as the lightness and contrast matches (Fig. 11). To be more exact, the mean brightness match, as measured in the low-articulated conditions, was 13.16.

#### Discussion

Our results are in line with those obtained by Zdravković, Economou, and Gilchrist (2006), who measured lightness of the same object divided by an illumination border. Specifically, their observers systematically judged the lightness of the object patch in highlight as lighter than that in shadow. Although they employed fewer (10) observers, the error bars in their graphs point out (especially for the blackest paper) some interindividual differences. The individual variations in lightness matches observed in

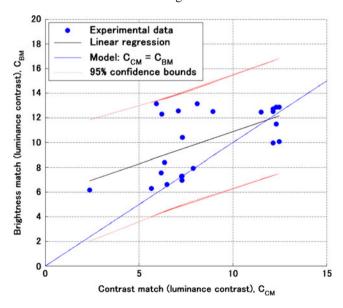
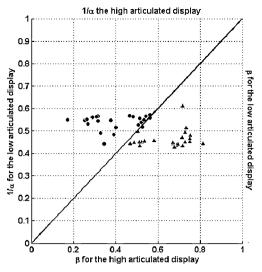


Fig. 10 Linear regression of mean brightness match on mean contrast match





**Fig. 11** Effect of articulation on brightness and contrast matches. This graph is similar to Fig. 8, except that the Brunswik ratios (triangles) were evaluated for brightness rather than lightness matches

our experiment were even larger. They corroborate the extraordinary range of individual differences observed in early studies on lightness (Katz, 1935; Thouless, 1932; Woodworth, 1938, pp. 601–607). The range of individual variations is so large that these do not seem to be just interindividual differences. There must be some factor causing such a large variation between observers' judgements in lightness constancy experiments.

In the present work, we employed three different tasks typical of experiments on lightness perception: lightness, brightness, and contrast matching (e.g., Gilchrist, 2006; Jameson & Hurvich, 1989). In the previous studies, the matches were found to depend considerably on the task (e.g., Arend & Spehar, 1993). It is generally believed that these tasks instruct observers on which perceptual dimension to base their judgements: "apparent reflectance" (lightness), "apparent luminance" (brightness), or contrast. Although our instructions were practically similar to those used by Arend and Spehar,9 whose results were different for all three instruction sets, we found no statistically significant difference between the lightness and contrast tasks, and minor, though significant, differences between the lightness and brightness tasks. Hence, in spite of the obvious difference between the lightness and contrast tasks, observers exhibited the same performance in both. Surprisingly, the brightness matches and

<sup>&</sup>lt;sup>9</sup> Their lightness-matching instruction was "to make the test patch 'look as if it were cut from the same piece of paper' as was the corresponding patch in the standard." The brightness-matching instruction was "to make the test patch 'have the same brightness as the corresponding patch in the standard, disregarding, as much as possible, other areas of the display. That is, make the amount of light coming from the test patch look the same as that from the standard'" (Arend & Spehar, 1993, p.450).

lightness matches were found to be strongly correlated. This means that although observers attempted different performance in the brightness task, they hardly succeeded. The profiles of the matches in all three of the tasks were found to be very similar. It seems that in all of the tasks, our observers based their matching practically on the same apparent dimension. What dimension might it have been?

It is rather unlikely that the dimension was luminance contrast. If our observers performed luminance contrast matches, why were they so inaccurate? Indeed, even for the contrast-matching task, the mean contrast match was found to be only 35% of the luminance contrast of the illumination border. This is in line with previous finding (Logvinenko, 2005b) that the luminance contrast produced by the illumination edge is underestimated by a factor of 2.2 as compared to that produced by the reflectance edges. In contrast, Arend and Spehar (1993) found reasonably accurate performance for their observers in the contrast-matching task. We believe that the key difference between their and our experimental displays was that they displayed stimuli on a CRT monitor screen, whereas we employed real surfaces illuminated by real light sources.

At this point, we agree with Gibson (1979), who argued that pictures are dual objects: On the one hand, they are either sheets of paper contaminated with paint or selfluminous objects, like computer monitor screens; on the other, they render some other objects (pictorial objects). There is an obvious difference between, say, a pictorial snowball and its real prototype, and we are always aware of this difference. We believe that a difference of the same sort exists between the pictorial lightness of the depicted snowball and the lightness of its physical prototype. It was found, for example, that the simultaneous lightness contrast illusion was rather strong for a pictorial wall of blocks (Adelson, 1993; Logvinenko, 1999), yet it either completely disappeared (Logvinenko, Kane, & Ross, 2002) or was considerably reduced (Logvinenko & Kane, 2003) when implemented as a real wall. Moreover, using the same multidimensional scaling technique as Logvinenko and Maloney (2006), Logvinenko, Petrini, and Maloney (2008) found that pictorial achromatic colours in the Adelson snakelike pattern made a one-dimensional manifold. In other words, no surface-brightness dimension was found, despite a clear pictorial impression of areas of different illuminations. This shows that, firstly, surface brightness emerges only when real (not pictorial) objects are perceived. Secondly, surface brightness cannot be reduced to brightness (i.e., perceived luminance). Thirdly, results obtained for pictorial displays cannot always be generalised to natural perception (i.e., perception of real, three-dimensional things).

The dual nature of a monitor screen pictures allows a dual mode of perception: objective versus pictorial. As has been suggested by many authors before (e.g., Arend, 1994),

when a pictorial attitude (facilitated by the instruction on brightness match) is taken, observers judge brightness, whereas under an objective attitude (lightness instruction), they judge pictorial lightness. However, our results show that the dual mode of perception is hardly possible for real scenes. Because both authors took part in the present experiment as observers, we testify that it was very hard to compare the luminance contrast of the light reflected from a material edge and that from an illumination edge. It was only when both edges were of the same type (i.e., both either material or illumination edges) that a luminance contrast match became easy to perform. Moreover, in this case, a luminance contrast match could be done very accurately. For instance, when stimuli are presented on a monitor screen, the accuracy of luminance contrast matching was found to be remarkably high (see, e.g., Nachmias & Sansbury, 1974). Thus, it seems to be extremely difficult, if not impossible, to evaluate the luminance contrast of the light reflecting from real objects.

For analogous reasons, brightness can be eliminated as the dimension underlying the observers' judgements in our experiments. Even in the brightness-matching task, the judgments were as inaccurate as in the contrast-matching task. As a matter of fact, the brightness of the spotlight was underestimated relative to that of the target squares by a factor of 2.4. (To be more exact, 2.4 was the luminance ratio for the mean-brightness-match target and the highlighted region.) Furthermore, if in this task observers really matched brightness, why did the brightness matches exhibit such a strong dependence on articulation of the spotlight? Finally, the individual variations were found to be as large in brightness matches as in lightness matches. However, Katz wrote that "these individual differences in vision can appear only when we are dealing with a comparison of colours standing under different conditions of illumination" (Katz, 1935, p. 139). More specifically, the same observers who exhibited large variations in lightness judgements of two discs (one in high light, and the other in shadow) viewing them directly, were unanimous concerning their brightness when viewing the discs through the reduction screen (Katz, 1935, p. 135). Therefore, if our observers based their judgements on brightness, we would expect much less individual variation.

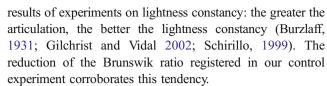
This raises a more general question: Are inexperienced human observers capable of evaluating the brightness of the reflected light in scenes constituted by real objects lit by real lights? The enormous individual variations and large inaccuracy in brightness matching observed in our experiment testify against such a capability. Yet, brightness matches become an easy task if a reduction screen is used, as in Katz's (1935) demonstration. However, the reduction screen destroys the object mode of perception. When using a reduction screen, we perceive a patch of light rather than a fragment of an object. Alternatively, adopting a special introspective attitude on pictorial vision, an experienced



observer can see what Gibson (1950) called the *visual field*, contrasting it to the *visual world* (i.e., the normal perceptual image of a real scene). Admittedly, it is possible to estimate the brightness of light in the visual field. However, the effect of the pictorial attitude is eventually similar to that of the reduction screen: It facilitates abandoning the object (normal) mode of perceiving. Nevertheless, our data testify that inexperienced observers cannot judge the brightness of the reflected light in the normal mode of perceiving real objects illuminated by real light.

Strictly speaking, exact lightness matches could not have been achieved in our experiment, either, because the lightness continuum was different on each side of the illumination border. As has been argued elsewhere (Logvinenko & Maloney, 2006), an asymmetric lightness match (i.e., a lightness match across an illumination border) is impossible in general, because of unavoidable difference in surface brightness. Logvinenko and Maloney suggested that, instead, observers could attempt to find the least dissimilar pair. In the context of the present study, this means that in response to the lightness-match instruction, observers might have tried to find in the shadowed region a target square least dissimilar from the highlighted part of the background. The problem with this hypothesis (referred to as the *least-dissimilarity* hypothesis) is that it predicts a perfect lightness match. Indeed, it was shown that when instructed directly to choose the least dissimilar papers lit by different lights, observers chose the papers with practically equal albedos (Logvinenko & Maloney, 2006). However, in the present study, lightness matches were far from perfect. Moreover, as in many previous lightness constancy studies, large interindividual differences were found. At first glance, the least-dissimilarity hypothesis is hardly compatible with the broad range of interindividual differences observed in lightness experiments.

There is, however, an important difference between our experimental display and that used by Logvinenko and Maloney (2006). In their experiment, the articulations of both the shadowed and the highlighted fields (i.e., the number of the patches with different albedo) were equal, whereas in our experiment, the spotlight was less articulated than the shadowed field. As mentioned in the introduction, articulation is an important feature that essentially affects the



Curiously, there is no explanation why articulation affects lightness perception. Moreover, there is no agreement even on what the very term *articulation* refers to (e.g., Maloney & Schirillo, 2002). Nor is there an explanation of the interindividual variations that are observed in many experiments on lightness constancy, including ours. We believe that the articulation effect and the interindividual variations are two serious challenges that any theory of lightness perception should take into account. Along with these, in the context of the present study we also face the following question: How could the least-dissimilarity hypothesis be reconciled with the effect of articulation and the interindividual variations?

We believe that the least-dissimilarity hypothesis can account for both interindividual differences and the effect of articulation if the two-dimensionality of achromatic colours is taken into consideration. When a scene is partitioned into frames of different illuminations, it is safe to assume (at least, to the first approximation) that, within a frame, lightness from the same lightness continuum is assigned according to the luminance ratio rule (Gilchrist, 2006). Patches in different frames are, generally, assigned lightness from different lightness continua. However, there is uncertainty concerning what lightness continuum should be chosen for what frame. Although this uncertainty can be partly resolved if the luminance contrast at the frame borders is assessed by the visual system, it always remains present in a scene (Logvinenko et al., 2008). For example, the same luminance contrast at the horizontal (illumination) border in Fig. 1 can be produced by a black paper lit by a bright light or a grey paper lit by a dim light. Generally, the same proximal stimulus (i.e., luminance contrast) can be produced by an infinite number of distal paper-light stimulus pairs. Note that given a particular value of luminance contrast, these pairs are constrained by the reciprocal relationship between paper albedo and light intensity. From the ecological point of view, one might expect this reciprocal relationship to have been interiorised for perception to be veridical. In other words, one might assume that a reciprocal relationship between lightness and surface brightness (in response to a particular proximal stimulus) is built into our perception as prior knowledge, as a heuristic reflecting the structure of the physical world.

An invariant relationship between apparent illumination and lightness was claimed by many early authors (for reviews, see Gilchrist, 2006, pp. 217–223; Logvinenko, 1997). Although there are many phenomenological demonstrations in favour of the trade-off between lightness and apparent illumination (e.g., Gilchrist, 1977; Logvinenko, 2005a;



The stimulus display in our experiment was presented against a black textureless background mounted on a dimly illuminated grey wall, whereas Logvinenko and Maloney (2006) used a white paper with a random-dot design as a background. Furthermore, the stimulus display in their experiment was inserted in a wooden box with a wooden divider placed on a table in an experimental room that was moderately illuminated. Thus, the difference in articulation was mainly due to the fact that Logvinenko and Maloney's experimental display was framed in the highly textured natural environment of the experimental room in comparison to which our stimulus situation was rather impoverished.

Logvinenko & Menshikova, 1994), direct experimental tests have yielded somewhat controversial results. Some of these findings support the apparent-illumination-versus-lightnessinvariance hypothesis (e.g., Kozaki & Noguchi, 1976; Logvinenko & Menshikova, 1994; Noguchi & Kozaki, 1985), whereas others have failed to confirm it (e.g., Beck, 1961; Kozaki, 1973; Oyama, 1968; Rutherford & Brainard, 2002). As pointed out elsewhere (Logvinenko, 1997, 2009), such a contradiction might have been caused, at least partly, by the inadequacy of the chosen measures of "apparent illumination." Furthermore, if, as we do believe, apparent illumination does not exist as an independent perceptual dimension on a par with lightness, all of the attempts to establish a quantitative relationship between elusive apparent illumination and lightness were doomed to fail from the outset.

The major difference between the apparent-illuminationversus-lightness-invariance hypothesis and what is referred to here as the reciprocal relationship (trade-off) between lightness and surface brightness comes from the ordinal nature of surface-brightness relations. More specifically, the trade-off between lightness and surface brightness implies, in particular, that the rectangular area of spotlight in Fig. 1 can be assigned either a blackish colour from some lightness continuum, or grey colour from another lightness continuum of lower surface brightness. Here, lower surface brightness is understood in the sense of the surface-brightness order defined above. If it is assigned blacker lightness, the lightness continuum should be of higher surface brightness. Note that we do not specify how much higher. Nor do we claim any exact quantitative relationship between lightness and surface brightness (as does the apparent-illumination-versus-lightness-invariance hypothesis). A qualitative reciprocity between lightness and surface brightness is enough to account for our, and many previous, results.

When lightness is assigned to the spotlight and background from the same lightness continuum, as happens in the case of the Gelb effect, the surface brightness of the shadowed part of the background and that of the highlighted part are equal, with their lightnesses being different. In this case, the border between these areas is perceived as a pure lightness border. In other words, here the Gelb effect takes place. When they are assigned the same lightness from different lightness continua, the border between the two areas is perceived as a pure surface-brightness border. In this case, we have lightness constancy. The Gelb effect and lightness constancy are the two extremes. There are an infinite number of intermediate cases when lightness values in the spotlight and in the

shadowed area are assigned from different lightness continua. In these cases, we will say that the *partial Gelb effect* takes place. Because "partial Gelb effect" seems to be a rather awkward word combination, we will use instead the term *partial illumination discounting*, simply as a synonym. (This logic, then, dictates using *illumination discounting* as synonymous with the Gelb effect.<sup>12</sup>)

Using more conventional terminology, one can say that in the case of partial illumination discounting, the illumination edge is perceived partly as a lightness (material) edge, and partly as a surface-brightness (illumination) edge. In our display (Fig. 1), partial illumination discounting would mean that the horizontal luminance border produced by the illumination edge is not perceived as a pure illumination edge. Instead, the border bifurcates into the lightness and surface-brightness edges, its luminance contrast splitting into lightness contrast and surface-brightness contrast. The proportion between lightness contrast and surface-brightness contrast indicates the strength of partial illumination discounting: the higher the lightness contrast at the illumination edge, the stronger the partial illumination discounting.

In this context, imperfectness of lightness constancy and partial illumination discounting are simply two sides of the same coin. If partial illumination discounting takes place, this means that some deviation from lightness constancy is observed, and vice versa. Therefore, the interindividual variability of lightness match observed in our experiment can be interpreted as that of partial illumination discounting. If we understand why different observers experienced partial illumination discounting of different strengths, we will understand what caused such enormous differences in lightness matches between individuals. We find the problem of interindividual variations much more tractable in terms of partial illumination discounting than in terms of lightness constancy. Likewise, we are convinced that it is easier to understand why articulation affects the magnitude of partial illumination discounting than why it would affect lightness

The major reason for the individual variations in partial illumination discounting is the principal ambiguity of

<sup>&</sup>lt;sup>11</sup> Note that lightness constancy does not mean that the colour appearance of the two areas is the same; there will be a difference in surface brightness. Because of this surface-brightness difference, an exact match between the two areas is impossible.

 $<sup>\</sup>overline{^{12}}$  Note that in contrast with the common usage of this term (i.e., as a hypothetical process), in this article it refers to a concrete visual phenomenon (i.e., the Gelb effect).

<sup>&</sup>lt;sup>13</sup> This does not mean, however, that we see a dual border in the sense that the surface-brightness edge overlaps with the lightness edge. In other words, we do not have what Gilchrist et al. (1983) called "two images in one." The border made by two achromatic colours from different lightness continua is as monistic as a border made by two shades of orange. In the latter case, we do not perceive a yellow border overlapped by a red border. Yet, for the sake of theoretical analysis, we can theoretically decompose the orange edge into the yellow and red edges. The same abstraction is being made here when we describe the achromatic colour border in terms of surface-brightness and lightness borders.

luminance contrast as a stimulus for achromatic colour. Even when lightness is considered as a one-dimensional continuum, the luminance-contrast ambiguity and a hypothetical process of resolving it are the major themes around which the theoretical discussion has been, implicitly or explicitly, revolving for the last two decades (for comprehensive reviews, see Gilchrist, 1994, 2006). Adopting the view of the two-dimensionality of achromatic colours makes the luminance contrast ambiguity problem even more profound. Indeed, apart from anchoring the luminance contrast value within a lightness continuum, one needs to decide which of many lightness continua to choose. In other words, luminance contrast should be anchored not only with respect to lightness, but also with respect to surface brightness. It should be noted, however, that these two anchoring processes are not independent of each other. The surface-brightness/lightness trade-off constrains anchoring with respect to lightness, provided that anchoring with respect to surface brightness has been accomplished. More specifically, a quantitative analysis shows that, if the luminance ratio rule (Wallach, 1963) holds true within an equiluminant frame, anchoring with respect to surface brightness predetermines anchoring with respect to lightness up to a multiplicative constant common to all frames (Logvinenko et al. 2008). This means that, practically, there is no need for local anchoring with respect to lightness once anchoring with respect to surface brightness has been accomplished. Thus, anchoring with respect to surface brightness is the key to anchoring with respect to lightness.

Anchoring with respect to surface brightness is intimately related to partitioning the whole scene into frames of equal surface brightness, which, in turn, relates to the so-called luminance edge classification (Gilchrist, 2006; Gilchrist, Delman, & Jacobsen, 1983). Since, by definition, the same anchor (with respect to surface brightness) is used within a frame, one can say that frames are what undergo surfacebrightness anchoring. The more frames there are in a scene, the fewer the degrees of freedom that exist for surfacebrightness anchoring. Therefore, scenes with multiple light sources (or with a complicated relief producing multiple reflexes) are less ambiguous in terms of surface-brightness anchoring. Note that our experimental display with just two frames (i.e., with one illumination border) is highly ambiguous in this sense (as have been most of the experimental displays used to investigate lightness constancy<sup>14</sup>).

<sup>&</sup>lt;sup>14</sup> It should be mentioned that in many previous experiments on lightness constancy (in which real surfaces and light sources, rather than pictures on computer monitors, were employed), there was no explicit illumination border as in our experimental display. Therefore, in these experiments the uncertainty in choosing lightness continua was even higher.



As we can see, lightness anchoring cannot be understood without surface-brightness anchoring, which, in turn, depends on partitioning into frames of equal surface brightness. This brings into focus the process of luminance edge classification. It is clear that classifying luminance edges into those produced by illumination and those produced by material has to be based on some optical and geometrical properties of the retinal luminance patterns that the visual system can use for this purpose (so-called illumination cues). Although the importance of edge classification for lightness perception was realised a few decades ago (Gilchrist et al. 1983), it has not been paid as much attention as it merits, perhaps because it has seemed that it would be possible to develop a theory of anchoring processes without resorting to this process (Gilchrist, 2006). Nevertheless, in the broader context of colour vision, a few illumination cues have been established (for a comprehensive review, see Kingdom, 2008). For instance, the straightness and fuzziness of a luminance edge were found to testify in favour of its being produced by an illumination edge; curvature, closedness, and sharpness, of its being produced by a material edge (Logvinenko, Adelson, Ross, & Somers, 2005). From this point of view, the rectangular spotlight in Fig. 1 is ambiguous. Indeed, on the one hand, it is sharp, curved, and closed; on the other, there are many other natural cues indicating in favour of an illumination border. 15 Given such ambiguity, it is hardly surprising that the strength of partial illumination discounting varied among observers.

It should be kept in mind, however, that even if there were no ambiguity at all, and all of the illumination cues confirmed the presence of an illumination edge, the uncertainty concerning the proportions into which to split the luminance contrast at the illumination border (between the lightness and surface-brightness contrasts) would remain; thus, partial illumination discounting could, generally, take place. Illumination cues can only signal the presence of an illumination border; they cannot specify the surface-brightness contrast at the border.

Articulation of the spotlight reduces this uncertainty, thus reducing illumination discounting, since it imposes constraints on the choice of the lightness continuum (Logvinenko, 2002). Indeed, the luminance ratio rule dictates that in the presence of the articulating squares, the spotlight background lightness must be blacker than that of all of the 12 articulating squares. Therefore, the lightness of the spotlight background should be blacker than the white present by at least 12 just-noticeable differences. The more articulating squares are present, the blacker the spotlight background will

<sup>&</sup>lt;sup>15</sup> If there were no such cues, the complete Gelb effect would have taken place. In other words, one would have seen no illumination edge at all. On the contrary, none of our observers had any doubts concerning the existence of the illumination border in our display.

appear. Hence, increasing the articulation of the spotlight should raise the Brunswik ratio. This is exactly what has been observed in our experiment. In the main experiment (high articulation), on average, observers showed considerably higher lightness constancy, as compared to the first control experiment (low articulation).

Thus, the imperfectness of lightness matches and the large variations in the Brunswik ratio observed in our experiment for different individuals can be accounted for by partial illumination discounting. More specifically, we argue that in the lightness-matching task, our observers chose the target square having the same lightness as (or the one closest to) the background lightness. However, because the lightness continuum from which lightness was assigned to the background was different due to partial illumination discounting, the albedo of the match square was different from that of the background. Because of the unavoidable uncertainty of partial illumination discounting, the lightness varied considerably from observer to observer.

As for the contrast-matching task, since the contrast matches of the whole group did not differ significantly from the lightness match, we believe that, in general, observers estimated lightness contrast in this task. However, for 27% of our observers, the contrast and lightness matches were significantly different. Therefore, these observers based their contrast matches not only on lightness contrast. Specifically, a combination of estimates of both the surface brightness and lightness contrasts at the illumination edge might have contributed to observers' contrast matches. Yet the resemblance of the matches in the two tasks indicates that the lightness contrast at the illumination border probably dominated the observers' decisions under the contrast-matching task.

It follows that illumination contrast discounting can be measured with the contrast-matching technique only in the absence of illumination discounting (i.e., when the illumination edge is perceived as purely a surface-brightness edge). It seems safe to suggest that in this case, observers would match the surface-brightness contrast at the horizontal illumination edge to the lightness contrast of the target squares. Note that only in this case (i.e., without illumination discounting), the contrast match can provide information for the illumination contrast discounting. When partial illumination discounting took place, then, observers could base their judgements, theoretically, on either surface-brightness contrast or lightness contrast, or on both. Therefore, the illumination contrast discounting index obtained in the present study cannot, strictly speaking, quantify illumination contrast discounting. This accounts for a considerable difference between illumination contrast discounting as measured with the contrast-matching tasks in the present study and the previous study (Logvinenko, 2005b), on the one hand, and with multidimensional scaling (Logvinenko & Maloney, 2006), on the other. This inconsistency could be resolved in an experiment with highly articulated displays (similar to those used by Logvinenko & Maloney, 2006) using both techniques.

It is worth mentioning that the results of our experiment for the brightness match task provide more evidence supporting the two-dimensional conception of achromatic colours that denies the apparent illumination as an independent perceptual dimension correlating with physical illumination (and against the classical view that apparent illumination and lightness are distinct and independent perceptual dimensions). Indeed, if our participants had access to an internal variable reflecting apparent illumination, then we would expect their brightness matches to correlate with the illumination of the highlighted area of the display. Furthermore, if their brightness matches rested upon the apparent illumination of the highlight, then, because apparent illumination is locked up with lightness in a reciprocal relation, their brightness matches should have reciprocally covaried with the lightness matches they produced. Instead, brightness matches were found to covary with lightness matches in the opposite direction (Fig. 10). There is only one explanation for this fact: participants do not have direct access to an internal variable reflecting apparent illumination. What information they have about illumination is indirectly inferred from the two-dimensional lightness manifold and is insufficient to enable accurate brightness matches. Therefore, as claimed in Introduction, surfacebrightness is not just a euphemism for apparent illumination.

The concept of the two-dimensional achromatic colour manifold and the phenomenon of illumination discounting prompts us to take a fresh look at some old controversial issues in lightness perception, such as, for example, the intrinsic-images-based approach, which suggests that the retinal illuminance pattern is analysed into two component images: an illumination image and a lightness image (see, e.g., Adelson & Pentland, 1996; Arend, 1994; Barrow & Tenenbaum, 1978; Bergstrom, 1977; Gilchrist, 1979). This approach implies, implicitly or explicitly, that each luminance edge is classified as either a lightness or an illumination edge. It follows from our analysis above that this approach predicts either perfect lightness constancy or a perfect Gelb effect. As this prediction is never obtained experimentally, the intrinsicimages-based approach has been fairly criticised (Gilchrist, 2006) and, finally, abandoned in favour of the anchoring theory (Gilchrist et al. 1999). This theory is based on a single lightness continuum and does not have room for apparent illumination at all. Actually, a step back has been taken to the times when the role of illumination in lightness perception was completely neglected. "A curious fact of current lightness theories is the rather unanimous position, either implicit or explicit, that illumination is not perceived" (Gilchrist et al. 1983, p.427).

Interestingly, the intrinsic-images-based approach itself does not require a none-or-all edge classification. Strictly

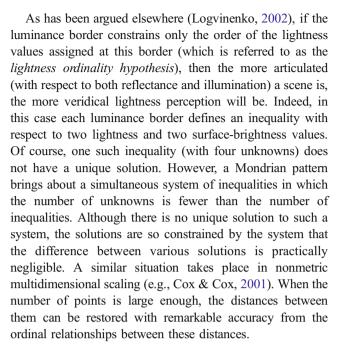


speaking, one can assume that a luminance border can be interpreted as partly a lightness and partly an illumination border. The concept of a two-dimensional achromatic colour manifold permits such an assumption without postulating apparent illumination as an independent perceptual dimension.

Furthermore, the intrinsic-images-based approach can be more flexible concerning luminance edge integration. Gilchrist's version of this approach implies linear integration, which entails a linear relationship between lightness and luminance. However, it is well established that the lightness-versus-luminance response function is not linear (e.g., Whittle, 1994; Wyszecki & Stiles, 1982, pp. 493–499). It was shown recently (Rudd, 2009) that a model of edge integration that takes into account a more plausible, nonlinear relationship between lightness and luminance accounts for the lightness range compression (Gilchrist et al. 1999). Therefore, taking into account nonlinear gain control reconciles the intrinsic-images-based approach with lightness range compression—the phenomenon that was one of the major reasons to abandon the intrinsic-images-based approach (Gilchrist, 2003, 2006).

In our experiment, lightness range compression revealed itself in that the highlighted background looked light grey rather than black (or very dark grey). Note that this is different from partial illumination discounting. Indeed, whatever the lightness continuum, the lightness range assigned to the articulating squares in the highlighted background seems to have been smaller than that predicted by the linear-edge-integration hypothesis. Note also that different observers exhibited lightness range compression of different magnitudes. This means that lightness range compression was found to be subject to large individual variations. It is highly unlikely that these variations could be explained by individual differences in contrast gain control.

The individual variations in lightness range compression observed in our experiments are in line with the hypothesis of an ordinal nature of the relationship between lightness and luminance ratio (Logvinenko, 2002). According to this hypothesis, the luminance ratios within a frame determine only lightness order rather than absolute lightness values. In other words, the luminance ratios within a frame specify not a unique series of shades of greyness, but an infinite number of such series. The ordinal nature of lightness accounts not only for numerous "errors" in lightness perception (especially in laboratory conditions), but also for the individual variations in these "lightness errors." However, one might ask: If there is no strict connection between luminance ratios and lightness, how can one explain veridical lightness perception (especially in the natural environment)? In other words, can we have a theory that accounts for both veridical and erroneous lightness perception? For example, Gibson (1979) did not believe that there can be a single theory for veridical and illusory perception.



Because the human visual system has evolved to perceive a natural environment that is highly articulated with respect to both reflectances and illuminants, this does not seem to be an unlikely way to encode lightness. Indeed, deriving lightness from the ordinal relationships at the luminance borders would secure veridical lightness perception in the natural environment. However, such a way of encoding lightness would be rather inefficient in the artificial conditions of vision laboratories when low-articulated stimulation is used.

### Conclusion

Measuring lightness constancy with an asymmetric lightnessmatching technique for nearly a century has brought about the following puzzling facts: Complete lightness constancy has been registered only for highly articulated stimulus displays. In most experimental studies, lightness constancy has been found to be imperfect, to depend on articulation, and to vary considerably across observers. We believe that all of this is a result of the inadequacy of asymmetric lightness matching for measuring lightness constancy. Because an asymmetric lightness match is impossible (Logvinenko & Maloney, 2006), we tried to find out what observers actually do when they are asked to make an asymmetric lightness match. Our results are compatible with the suggestion of Logvinenko and Maloney that when forced to make an asymmetric lightness match, observers choose the least dissimilar pair of differently illuminated papers. The deviation from complete lightness constancy, dependency on articulation, and interindividual differences can be accounted for by partial illumination discounting—that is, by perceiving a luminance border produced by an illumination edge as partly a surface-



brightness edge and partly a lightness edge. It was found that the stronger the illumination discounting, the lower the Brunswik ratio. Complete illumination discounting leads to complete lightness *inconstancy*. To see lightness constancy, one must have no illumination discounting at all. In other words, either the perceptions of material and illumination edges are both veridical, or both are not veridical.

**Author Note** This work was supported by EPSRC Grant EP/C010353/1 to A.D.L.

#### References

- Adelson, E. H. (1993). Perceptual organization and the judgement of brightness. Science, 262, 2042–2044.
- Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 409–423). New York: Cambridge University Press.
- Arend, L. (1994). Surface colours, illumination, and surface geometry: Intrinsic-image models of human color perception. In A. L. Gilchrist (Ed.), Lightness, brightness, and transparency (pp. 159–213). Hillsdale, NJ: Erlbaum.
- Arend, L. E., & Spehar, B. (1993). Lightness, brightness, and brightness contrast: 1. *Illuminance variation*. *Perception & Psychophysics*, 54, 446–456.
- Barrow, H. G., & Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. In A. R. Hanson & E. M. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press
- Beck, J. (1961). Judgments of surface illumination and lightness. *Journal of Experimental Psychology, 61,* 368–373.
- Bergstrom, S. S. (1977). Common and relative components of reflected light as information about the illumination, colour, and three-dimensional form of objects. Scandinavian Journal of Psychology, 18, 180–186.
- Burzlaff, W. (1931). Methodologische Beitrage zum Problem der Farbenkonstanz [Methodological notes on the problem of color constancy]. Zeitschrift für Psychologie, 119, 117–235.
- Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press. Cox, T. F., & Cox, M. A. A. (2001). *Multidimensional scaling* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Craven, B. J., & Foster, D. H. (1992). An operational approach to colour constancy. *Vision Research*, 32, 1359–1366.
- Foster, D. H. (2003). Does colour constancy exist? Trends in Cognitive Sciences, 7, 439–443.
- Foster, D. H., Nascimento, S. M. C., Amano, K., Arend, L., Linnell, K. J., Nieves, J. L., Foster, J. S. (2001). Parallel detection of violations of color constancy. *Proceedings of the National Academy of Sciences*, 98, 8151–8156.
- Gelb, A. (1929). Die Farbenkonstanz der Sehdinge. In A. Bethe, G. von Bergmann, G. Embden, & A. Ellinger (Eds.), Handbuch der normalen und pathologischen Physiologie (pp. 594–678). Berlin: Springer.
- Gibson, J. J. (1950). The perception of the visual world. Cambridge, MA: Riverside Press.
- Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.
- Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. Science, 195, 185–187.
- Gilchrist, A. (1979). The perception of surface blacks and whites. Scientific American, 240, 112–123.

- Gilchrist, A. (1994). Absolute versus relative theories of lightness perception. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 1–33). Hillsdale, NJ: Erlbaum.
- Gilchrist, A. (2003). The importance of errors in perception. In R. Mausfeld & D. Heyer (Eds.), Colour perception: Mind and the physical world (pp. 435–451). Oxford: Oxford University Press.
- Gilchrist, A. (2006). Seeing black and white. New York: Oxford University Press.
- Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and interpretation of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, 33, 425–436.
- Gilchrist, A., & Economou, E. (2003). Dualistic versus monistic accounts of lightness perception. In L. Harris & M. Jenkin (Eds.), Levels of perception (pp. 11–22). New York: Springer.
- 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.
- Gilchrist, A., & Vidal, A., Jr. (2002). Articulation effects in lightness: Historical background and theoretical implications. *Perception*, 32, 141–150.
- Hering, E. (1964). *Outlines of a theory of the light sense*. Cambridge, MA: Harvard University Press.
- Jameson, D., & Hurvich, L. M. (1989). Essay concerning color constancy. Annual Reviews of Psychology, 40, 1–22.
- Katz, D. (1935). *The world of colour.* London: Kegan Paul, Trench, Trubner, & Co.
- Kingdom, F. A. A. (2008). Perceiving light versus material [Invited review]. Vision Research, 48, 2090–2105.
- Koffka, K. (1935). Principles of Gestalt psychology. New York: Harcourt, Brace.
- Kozaki, A. (1973). Perception of lightness and brightness of achromatic surface color and impression of illumination. *Japa-nese Psychological Research*, 15, 194–203.
- Kozaki, A., & Noguchi, K. (1976). The relationship between perceived surface-lightness and perceived illumination. *Psychological Research*, 39, 1–16.
- Krantz, D. M., Luce, R. D., Suppes, P., & Tversky, A. (eds.) (1971).
  Foundations of measurement: vol. 1. Additive and polynomial representations. New York and London: Academic Press.
- Logvinenko, A. D. (1997). Invariant relationship between achromatic colour, apparent illumination, and shape of surface: Implications for the colour perception theories. In C. M. Dickinson, I. J. Murray, & D. Carden (Eds.), *John Dalton's colour vision legacy* (pp. 591–597). London: Taylor & Francis.
- Logvinenko, A. D. (1999). Lightness induction revisited. *Perception*, 28, 803–816.
- Logvinenko, A. D. (2002). Articulation in the context of edge classification. *Perception*, 31, 201–207.
- Logvinenko, A. D. (2005a). Does luminance contrast determine lightness? *Spatial Vision*, *18*, 337–346.
- Logvinenko, A. D. (2005b). On achromatic colour appearance. In J. L. Nieves & J. Hernánderez-Andrés (Eds.), AIC Colour '05: Proceedings of the 10th Congress of the International Colour Association, Part 1 (pp. 639–642). Granada: International Colour Association.
- Logvinenko, A. D. (2009). Pseudoscopic colour illusions. In D. Smith, P. Green-Armytage, M. A. Pope, & N. Harkness (Eds.), Proceedings of the 11th Congress of the International Colour Association. Colour Society of Australia: Sydney.
- Logvinenko, A. D., Adelson, E. H., Ross, D. A., & Somers, D. (2005). Straightness as a cue for luminance edge classification. *Perception & Psychophysics*, 67, 120–128.
- Logvinenko, A. D., & Kane, J. (2003). Luminance gradient can break background-independent lightness constancy. *Perception*, 32, 263–268.
- Logvinenko, A. D., Kane, J., & Ross, D. A. (2002). Is lightness induction a pictorial illusion? *Perception*, 31, 73–82.



- Logvinenko, A. D., & Maloney, L. T. (2006). The proximity structure of achromatic colours and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, 68, 76–83.
- Logvinenko, A. D., & Menshikova, G. Y. (1994). Trade-off between achromatic colour and perceived illumination as revealed by the use of pseudoscopic inversion of apparent depth. *Perception*, 23, 1007–1023.
- Logvinenko, A. D., Petrini, K., & Maloney, L. T. (2008). A scaling analysis of the snake lightness illusion. *Perception & Psycho*physics. 70. 828–840.
- MacLeod, R. B. (1940). Brightness constancy in unrecognized shadows. *Journal of Experimental Psychology*, 27, 1–22.
- Maloney, L. T., & Schirillo, J. A. (2002). Colour constancy, lightness constancy, and the articulation hypothesis. *Perception*, 31, 135–139.
- Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research*, 14, 1039–1042.
- Noguchi, K., & Kozaki, A. (1985). Perceptual scission of surfacelightness and illumination: An examination of the Gelb effect. *Psychological Research*, 47, 19–25.
- Oyama, T. (1968). Stimulus determination of brightness constancy and the perception of illumination. *Japanese Psychological Research*, 10, 146–155.
- Rudd, M. E. (2009). An edge-based account of lightness compression and insulation in the staircase Gelb effect [Abstract]. *Journal of Vision*, 9(8), 361a. doi:10.1167/9.8.361
- Rutherford, M. D., & Brainard, D. H. (2002). Lightness constancy: A direct test of the illumination estimation hypothesis. *Psychological Science*, 13, 142–149.

- Schirillo, J. A. (1999). Surround articulation: II. Lightness judgments. *Journal of the Optical Society of America A, 16,* 804–811.
- Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research*, 3, 263–346.
- Thouless, R. H. (1932). Individual differences in phenomenal regression. *British Journal of Psychology*, 22, 216–241.
- Tokunaga, R., & Logvinenko, A. D. (2010a). Hue manifold. *Journal of the Optical Society of America A*, 27, 2551–2557.
- Tokunaga, R., & Logvinenko, A. D. (2010b). Material and lighting dimensions of object colour. Vision Research, 50, 1740–1747.
- Tokunaga, R., & Logvinenko, A. D. (2010c). Material and lighting hues of object colour. Ophthalmic and Physiological Optics, 30, 611–617.
- Tokunaga, R., Logvinenko, A. D., & Maloney, L. T. (2008). Dissimilarity of yellow-blue surfaces under neutral light sources differing in intensity: Separate contributions of light intensity and chroma. Visual Neuroscience, 25, 395–398.
- Wallach, H. (1963). The perception of neutral colours. Scientific American, 208, 107–116.
- Whittle, P. (1994). Contrast brightness and ordinary seeing. In A. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 111–157). Hillsdale: Erlbaum.
- Woodworth, R. S. (1938). Experimental psychology. New York: Holt. Wyszecki, G., & Stiles, W. S. (1982). Color science: Concepts and methods, quantitative data and formulae (2nd ed.). New York: John Wiley & Sons.
- Zdravković, S., Economou, E., & Gilchrist, A. (2006). Lightness of an object under two illumination levels. *Perception*, 35, 1185–1201.

