Lightness and brightness judgments of coplanar retinally noncontiguous surfaces

James A. Schirillo and Steven K. Shevell
Department of Psychology and Department of Ophthalmology and Visual Sciences, University of Chicago, Chicago, Illinois 60637

Received November 16, 1992; revised manuscript received May 24, 1993; accepted June 2, 1993

Several experiments reveal that judgments of lightness and brightness of an achromatic surface depend, in part, on the luminances of other surfaces perceived to share the same depth plane, even if the surfaces are well separated on the retina. Two Mondrians, simulated on a CRT, were viewed through a haploscope. The more highly illuminated Mondrian contained a comparison patch and appeared nearer than the more dimly illuminated Mondrian, which contained the test patch. By independently varying the disparity of the test patch, observers could make the test patch appear to be in the depth plane of either the dimly or the highly illuminated Mondrian. Observers set the luminance of the test patch to match that of the comparison patch. The test was set as high as 15% more luminous when it was perceived in the depth plane of the highly illuminated rather than the dimly illuminated Mondrian. Both brightness and lightness judgments were affected by the perceived depth of the test, although the lightness judgments of inexperienced observers sometimes were dominated by local-contrast matching.

INTRODUCTION

In nature, surfaces sharing a common depth plane often share the same illumination, while regions in different depth planes often differ in illumination. Because coplanar surfaces may not be retinally adjacent to one another, the illumination falling on retinally contiguous surfaces may differ. Consequently, the brightness of a surface in a three-dimensional scene may depend on more than local luminance contrast and on more than simply incorporating the luminances of distant regions irrespective of their perceived depth. If judging the perceived brightness of a surface requires an inference about the light illuminating it, then the three-dimensional relations among objects may be an important factor.

If coplanar regions share the same illumination, then one possible way to assess the illumination falling onto a surface is perceptually to group surfaces that share a common depth plane, regardless of where light from these surfaces strikes the retina. The coplanar illumination hypothesis specifies that the perceived brightness of a surface may vary with the perceived depth of the surface. For example, imagine judging the brightness of a gray shirt as it hangs in a dimly lit closet surrounded by other clothes. Then imagine judging its brightness as it is moved into a well-lit room while it is still "retinally" surrounded by the clothes hanging in the closet. In the latter case the luminance of only the shirt, not the clothing that surrounds it, is increased. Attributing the increase in luminance to additional illumination falling on the shirt will increase the perceived brightness of the shirt. This would make the shirt's perceived brightness correlate with the illumination of the depth plane in which it lies. If, however, the increased difference in luminance is not attributed to a higher level of illumination, the perceived lightness of the shirt will increase; that is, the shirt will appear whiter instead of brighter. Lightness and brightness are separate perceptual dimensions. Lightness is judged on a relative scale: a surface is perceived as either lighter or darker than a fixed standard in a scale of grays. Brightness is judged on an absolute scale: a surface is perceived as brighter or dimmer, more akin to perceived intensity.

The coplanar illumination hypothesis is an underpinning of other models of lightness and brightness perception; it provides essential groundwork for Gilchrist's coplanar ratio hypothesis. Gilchrist argues that the perceived lightness of a surface depends on its contrast with apparently coplanar surfaces, while noncoplanar surfaces have virtually no effect. The coplanar ratio hypothesis provides a means by which "perceived lightness might be determined primarily by ratios within perceived planes rather than by all retinal ratios regardless of perceived depth" (Ref. 2, p. 186; see also Ref. 11).

In the current study, observers judged the perceived brightness and lightness of an achromatic surface by adjusting the luminance of a test patch to match a comparison patch. The test patch was varied in perceived depth. An important feature of the experimental design is that we assessed inferred illumination without varying retinal illumination. This is unlike previous studies of adaptation and constancy, in which the experimenters varied the illuminating light to determine the degree to which it was discounted. In that we varied only stereo disparity, the significance of altering the three-dimensional representation was isolated while keeping the two-dimensional retinal image virtually unchanged.

Overall, lightness and brightness judgments covaried in the predicted direction with the perceived depth of the test patch. The lightness judgments were in the direction of constancy. However, the magnitude of the lightness and brightness effects was much smaller than...
Fig. 1. Diagram of the stimulus pattern on the CRT screen. The entire upper Mondrian and lower test patch had 34' of crossed retinal disparity. Each Mondrian had a 6:1 range of simulated surface reflectance. The luminance of each patch in the lower Mondrians was one fifth that of its corresponding patch in the upper Mondrians.

predicted if only the coplanar illumination was considered. The results indicate that the inferred illuminant depends only in part on coplanar surfaces.

METHODS

Observers
Three observers with normal acuity and normal stereo-acuity were tested. Author JS, a 35-year-old male, was knowledgeable about the experimental paradigm and had experience in making brightness and lightness matches, using Mondrians in two-dimensional displays. SG, a 24-year-old female graduate student, had experience with simple three-dimensional chromatic displays but not with achromatic Mondrians and was naive regarding the experimental paradigm. BP, a 19-year-old female undergraduate, was an inexperienced observer who was also naive about the purpose of the experiment.

Apparatus
Gray patterns were generated with a Pixar II image processor under the control of a Sun 3/150 work station and were presented on an accurately calibrated Sony 17-in. (43-cm) color monitor. The 1280 × 1024-pixel screen was set to provide a steady, neutral background with an average luminance of 17.4 cd/m² and CIE chromaticity x = 0.33 and y = 0.33. The scan rate was 60 Hz noninterlaced. The red, green, and blue guns were linearized by use of a 9-bit look-up table. A given chromaticity and luminance, set by software, did not vary appreciably over the effective viewing area. The luminance was approximately constant (±3%) within the central region of the screen that displayed the four Mondrians.

The monitor was viewed haploscopically at a distance of 40 in. (100 cm) in a dark room. Two 45° mirrors, one directly in front of each eye, could be varied in distance from the eyes so that the observer always perceived a crisp three-dimensional image. The left-hand half of the CRT screen projected an image to only the left eye, while the right-hand half of the CRT screen projected an image to only the right eye.

Stimuli
The subjects perceived a three-dimensional image by viewing the CRT screen through the haploscope (Fig. 1). The CRT displayed four Mondrians simultaneously, each 5° high × 6.2° wide. Identical Mondrians located in the upper-left-hand and upper-right-hand quadrants of the CRT screen had 34' of crossed retinal disparity. Two other Mondrians, identical to each other, were located in the lower-left-hand and lower-right-hand quadrants of the CRT screen and had no retinal disparity. When the left and right Mondrians binocularly fused with use of the haploscope, the upper (comparison) Mondrian appeared to be closer to the observer than the lower (test) Mondrian (Fig. 2).

The entire image was composed of lights covering a 30:1 luminance range. The comparison (i.e., upper and nearer) Mondrian was composed of 12 patches that varied in luminance over a 6:1 range. These patches simulated a 6:1 range of surface reflectance. Twelve patches of identical shape and location formed the test (i.e., lower and farther) Mondrian. The luminance of each patch in the test Mondrian was set at one fifth of the luminance of its corresponding patch in the comparison Mondrian. This fivefold difference in luminance created an overall appearance of two Mondrians having identical surface reflectances, with each under a distinct illumination. Grouping surfaces of higher luminance within the comparison Mondrian made it appear brightly illuminated (we refer to this as 100% illumination). Grouping surfaces of lower luminance within the test Mondrian made it appear dimly illuminated (we refer to this as 20% illumination).

The upper Mondrian contained at its center a 1° × 1° comparison patch, the luminance of which was set by the experimenter. The simulated reflectance was 36%, 51%, 68%, or 90% (Munsell value v = 6.5/, 7.5/, 8.5/, or 9.5/, respectively).
The comparison patch was centered within a 3° × 3°-square area with a simulated reflectance of 25% (Munsell value \( v = 5.5 \)). The other patches within each Mondrian ranged in simulated reflectance from 5% to 92% (Munsell value \( v \) from 2.5/ to 9.6/). The geometric mean reflectance across the entire display (both dimly and brightly illuminated Mondrians) was 29.5% (Munsell value \( v = 6.0/; 17.4 \text{ cd/m}^2 \)).

The test-patch luminance was varied by the observer. In the no-depth condition a 1° × 1° test patch, centered within a 3° × 3° surround and located at the center of the lower Mondrian, had no disparity. The test patch appeared to lie in the same depth plane as its retinally adjacent surround and the remainder of the test Mondrian. In the depth condition the test patch had 34' of crossed retinal disparity. It appeared to lie in the same depth plane as the upper, nearer, retinally nonadjacent comparison Mondrian (see Fig. 2). The small shift in retinal disparity kept the 1° × 1° test patch close to the center of each eye's 3° × 3° surround. Thus retinal contrast was virtually the same in the no-depth and depth conditions. In the depth condition, the "nearer" test was perceived to fall in the center of the "farther" 3° × 3° surround.

Procedure
Observers participated in several practice sessions to ensure that they understood the instructions. Head position was maintained with a chin rest. Observers dark adapted for 3 min and then light adapted for 3 min to a uniform field at the average level of the Mondrian luminances. Adaptation was followed by observation of the Mondrian images described above. Observers easily and immediately fused the Mondrians so that they appeared three dimensional. An experimental session consisted of 20 depth trials randomly intermixed with 20 no-depth trials. The comparison patch was one of four different luminances selected randomly from trial to trial. On the first trial for each comparison luminance, the initial level of the test was set equal to its immediate surround.

Observers used a method of adjustment to set the test luminance so that the test and comparison patches matched perceptually. They controlled the luminance of the test patch by pressing separate buttons on a three-button computer mouse. One button incremented test luminance; another button decremented test luminance. The third button signaled that a satisfactory match had been achieved, at which point the test level was recorded and the trial ended. Between trials the CRT screen was uniform gray for 3 s. Then the next trial began. On subsequent trials the initial test luminance was randomly offset ±10% from the previous setting for the given comparison-patch luminance. Each session took approximately 1 h.

Instructions to subjects were as follows:
Test patch adjustment: "You will be asked to adjust a single test patch in the lower half of the display so that it matches its corresponding comparison patch in the upper half of the display. Throughout the experiment you are to spend about the same amount of time looking at the upper and lower halves of the display by alternating your gaze between them about once every two seconds."

Brightness judgments: "Using the mouse in front of you, vary the intensity of the lower test patch until it has the same brightness as the designated comparison patch in the upper half of the display. Throughout the experiment you are to spend about the same amount of time looking at the upper and lower halves of the display by alternating your gaze between them about once every two seconds."

Lightness judgments: "Using the mouse in front of you, vary the intensity of the lower test patch until it has the same lightness as the designated comparison patch in the upper half of the display. Throughout the experiment you are to spend about the same amount of time looking at the upper and lower halves of the display by alternating your gaze between them about once every two seconds."

RESULTS
Tests of a Coplanar Illumination Hypothesis
Experiment 1: Brightness Matches with Gray Surrounding the Test and the Comparison
The brightness of a test patch on a dimly illuminated Mondrian was matched to a comparison patch on a brightly illuminated Mondrian. Both the test and the comparison patches were centered within a gray surround of the same simulated reflectance.

The test-patch luminance for a brightness match (symbols, Fig. 3) was always set below the physical luminance...
of the comparison patch (dashed lines). Figure 3 shows that observers judged the test patch within the dimly illuminated Mondrian to be too bright when it was physically equal to the comparison patch within the brightly illuminated Mondrian. Observers thus reduced the test luminance away from a physical-luminance match.

This is expected from classical brightness contrast (however, see Experiment 9 below). The more important result, however, is that the perceived depth of the test patch affected the brightness measurements (circles above squares, Fig. 3). Observers set the test luminance, on average, 16% higher when it was perceived in the near, highly illuminated depth plane than when it was perceived in the dimly illuminated plane (JS, \( p < 0.001 \); SG, \( p < 0.001 \); BP, \( p < 0.05 \)). This difference is in the expected direction if the observer is inferring that a test coplanar with the highly illuminated comparison Mondrian is under a higher level of illumination than a test in the plane of a dimly illuminated Mondrian. This difference suggests also that retinal contrast, which remains constant across depth conditions, is not solely responsible for removing the effect of varying illumination.

Although information from coplanar surfaces affects the brightness match in the direction expected for inferred illumination, the effect of moving the test to the near-depth plane is quantitatively much less than of actually inferring a fivefold increase in the illumination falling on the test. If observers had completely accounted for the fivefold difference in luminance between the test and comparison Mondrians, they would have increased the test level by 500% in the near-depth plane.

**Experiment 2: Lightness Matches with Gray Surrounding the Test and the Comparison**

If observers infer a higher illumination when the test is perceived in the plane of the nearer, more highly illuminated Mondrian, then lightness judgments also should show an increase in the luminance of the test when the test appears in depth. Observers made lightness judgments (symbols, Fig. 4) using the same stimuli that they used for brightness judgments. As expected, the lightness measurements differed markedly from the brightness settings (symbols, Fig. 3). They show approximate lightness constancy (dotted lines) when the test remained within the depth plane of the dimly illuminated Mondrian (squares, Fig. 4). Only experienced observer JS, however, set the test to a significantly higher luminance when it was perceived in the highly illuminated depth plane (\( p < 0.001 \)). Under a higher perceived illumination, an increase in luminance is required for maintaining an in-depth test of constant reflectance. On the other hand, the perceived depth of the test caused no significant difference in the lightness matches of the other observers (SG, \( p > 0.25 \); BP, \( p > 0.50 \)).

**Experiment 3: Black Surrounding the Test and the Comparison**

The lightness results in Experiment 2 showed clear differences among observers. We considered whether the inexperienced observers, SG and BP, might have interpreted the lightness instructions as asking for simple local-contrast matches. If so, luminance contrast at retinally adjacent edges would dominate the lightness computation. In order to determine whether this was the case, we remeasured the effect of depth on lightness and brightness with a 3" x 3" black patch immediately surrounding both the test and the comparison (simulated reflectance of 0.0%; Munsell value \( v = 0/0 \)). This made the surrounding patches in each plane equal to each other and eliminated the possibility of judgments based on relative gray scales of adjacent patches (the surround's luminance is not affected by the level of illumination, because its reflectance is 0%). The geometric mean reflectance across both Mondrians was now 13.0% (Munsell value \( v = 4.2/7.7 \text{ cd/m}^2 \)).

As in the first experiment, varying the perceived depth of the test caused a luminance difference for brightness matches (filled circles above filled squares, Fig. 5). On average, the test was made 8% more luminous when it was perceived in the highly illuminated depth plane (JS, \( p < 0.001 \); SG, \( p < 0.02 \); BP, \( p < 0.05 \)). This is half of the magnitude obtained with the gray surrounds. With black surrounds the brightness matches fell closer to a physical-luminance match (filled circles and filled squares are nearer to the dashed line than in Fig. 3). This effect is expressed by the slope of the brightness results, which increases, on average, from 0.64 to 0.82 when the simulated surround reflectance is reduced from 25 to 0%. The
slope for physical-luminance matching is 1.0 and for perfect-reflectance matching is 0.20.

More important is that the lightness judgments for all three observers now show a clear effect of varying the depth plane of the test. On average, the test luminance was set 17% higher (open circles above open squares; Fig. 5) when it was perceived in the more highly illuminated depth plane (JS, p < 0.001; SG, p < 0.01; BP, p < 0.001). However, as with brightness judgments on either gray or black surrounds, the magnitude of change in the lightness settings with test depth falls far short of a physical-luminance match. This implies that, while observers' brightness and lightness judgments require a higher test luminance in the depth condition, the illumination of the near, highly illuminated Mondrian is not actually inferred to be the illumination that is falling upon the test.

Evidence with Test, Surround, or Mondrian Depth-Plane Variations

Experiment 4: Test at Various Depths

The above results suggest that the inferred test illumination is influenced by its coplanar Mondrian. Next we consider whether observers infer that the test is falling under increasing amounts of the near Mondrian's higher illuminant \( \text{illuminant}^{180} \) as the test disparity is gradually increased from zero. The perceived illumination would reach a maximum when test and Mondrian appear coplanar. Additional increases in disparity would make the test appear closer to the observer than the near Mondrian but should not alter the test's perceived illumination. This rationale implies that the Mondrian surfaces provide information that observers use in their inference of illumination throughout the three-dimensional space that the surfaces enclose.

The original experiment was repeated, except that the test disparity was varied gradually. The test was presented at either 0% depth (the previous no-depth condition) or 100% depth (the previous depth condition) or at 20%, 50%, or 80% of the disparity of the previous depth condition. In a series of separate runs the test disparity was increased further, appearing at either 100% depth (the previous depth condition) or 200% depth (twice the disparity of the previous depth condition) or at 120%, 150%, or 180% of the disparity of the previous depth condition.

When the observer was making brightness judgments, the test luminance set by the observer increased as the test depth increased to 100% (coplanar with near Mondrian). Figure 6(a) shows for observer JS the per-
percentage change in test luminance that is due to the depth of the test (i.e., luminance with depth minus luminance with no depth, divided by the luminance with no depth). Each symbol shape represents a different comparison luminance. Each dashed curve is a two-parameter fit with a linear increase to 100% disparity and no further increase thereafter. The test at 0% depth (no depth) differs significantly from those at 20% (p < 0.01), 50% (p < 0.001), 80% (p < 0.001), and 100% (p < 0.001). Also, the four nonzero test depths significantly differ from one another (p < 0.05), except that 80% depth does not differ from 100% depth. There is no significant increase in test luminance as the test depth is moved closer than the near Mondrian, i.e., from 100% to 200% (p > 0.50). The smallest comparison luminance (0.36) produced the largest relative change in test luminance that was due to test depth [diamonds, Fig. 6(a)]

Overall, observer SG's brightness judgments showed a similar though smaller depth effect [Fig. 6(b)]. The test depth at 0% (no depth) differs significantly from 50% (p < 0.01), 80% (p < 0.05), and 100% (p < 0.001); and the 20% and 50% tests differed significantly from the 80% and 100% tests (p < 0.01). There is no increase in test luminance as the test depth is moved closer than the near Mondrian, i.e., from 100% to 200% (p > 0.50).

Experiment 5: Depth of Test and Surround Covary
We next considered whether local edge contrast depends on inferred illumination. One possibility is that perceived local contrast depends only on retinal stimulation and not on inferred illumination. On the other hand, if illumination is inferred first, then perceived local contrast will vary when the perceived location of the test is moved to the near, more highly illuminated depth plane, because the test's immediate surround remains in the dim, far depth plane. In the current experiment the perceived depth plane of the test and its surround are varied together; thus they are kept under the same inferred illumination even as the test varies in depth.

The luminance of the 3\times3 gray surround contiguous with the test was held fixed at its previous luminance [reflectance of 25%, Munsell value v = 5.5/, dim (20%) illumination] but was varied in disparity with the test so that it could be perceived in the depth plane of the highly or dimly illuminated Mondrian. The disparity of the test and its surround always varied together. When the test and its surround appeared in the depth plane of the near, highly illuminated Mondrian, the observers set brightness matches that were 15% more luminous, on average, than those made in the depth plane of the far, dimly illuminated Mondrian (symbols, Fig. 7; JS, p < 0.001; SG, p < 0.001). These results are virtually the same as in the first experiment, in which only the test patch was varied in depth (symbols from Fig. 3 are replotted as dashed lines in Fig. 7). These results demonstrate, again, that the perceived brightness of the test patch, with or without a coplanar surround, is not determined solely by a matching of the local perceived edge contrast.

Moreover, moving the surround and the test to a more highly illuminated depth plane maintained the change in brightness seen when the test alone was changed in perceived depth. This suggests that inferred illumination affects brightness at a stage that follows perceived local contrast.

Experiment 6: Mondrians in Two Dimensions
It has been reported that changes in test disparity modify the effects of local neural retinal connections, such as lateral inhibition, thereby altering lightness and brightness judgments. We show here that coplanar surfaces, not simply the perceived depth plane of only the test patch, are the critical factors that mediate the depth effect. To determine the effect of merely separating in perceived depth the test and its surround, we placed both Mondrians of Experiment 1 in the same (far) depth plane, and only the perceived depth of the test patch was varied. This procedure preserved each Mondrian's separate illumination while removing depth as a basis for inferring illumination.

Overall, test depth had no effect on brightness judgments (symbols, Fig. 8; JS, p > 0.15; SG, p > 0.35; BP, p > 0.50). There was at most a possible trend at the lowest level of increment (0.36) toward slightly higher luminance levels with increasing depth of test.
Experiment 7: Test behind Far Mondrian

In this experiment the test was positioned in depth behind the far Mondrian, at a disparity equivalent in magnitude to that of the test when it had been coplanar with the nearer Mondrian in the original experiment.\textsuperscript{25} Gogel and Mershon\textsuperscript{2} and Mershon and Gogel\textsuperscript{27} would predict the same degree of neural uncoupling from the surround whether the test is in front of or behind the farther Mondrian.\textsuperscript{28}

Observers judged brightness in the current experiment by setting the test to a lower luminance (Fig. 9, filled circles below filled squares, by 8% on average) when it was perceived to be behind the far Mondrian (JS, $p < 0.001$; SG, $p < 0.001$; BP, $p < 0.05$).

These findings are in the direction opposite to that of those obtained in Experiment 1 (see Fig. 3) and therefore do not support Gogel and Mershon's hypothesis of correlating perceived stereo depth with neural uncoupling. Instead, they suggest that observers infer an even lower illuminant than that of the far Mondrian when the test is perceived to lie behind it. This inference may be due to the black background of the screen that extends beyond the Mondrians.

Experiment 8: Comparison behind Near Mondrian

From previous experiments we have argued that the brightness of a patch is due in part to an illuminant inferred from other surfaces in the same depth plane. If this interpretation is correct, then changing the perceived depth plane of the comparison to that of the far Mondrian should increase the perceived brightness of the comparison patch (which has a fixed luminance). This hypothesis was tested with a test field that remained coplanar with the far Mondrian. Unlike in the previous experiments, the comparison field was varied in depth so that it was coplanar with either its usual near Mondrian or with the dimly illuminated (far) Mondrian. This makes the in-depth position of the comparison analogous to the in-depth position of the test patch in the first experiment.

Brightness judgments resembled those obtained in the original experiment (symbols, Fig. 10). Observers set the test, on average, 10% more luminous when the comparison was perceived in depth (i.e., within the dimly illuminated far plane) than when the comparison was perceived in no depth (i.e., the brightly illuminated near plane) (JS, $p < 0.01$; SG, $p < 0.02$; BP, $p < 0.005$).

Evidence with Illumination Variations

Experiment 9: Mondrians with Ungrouped Luminances

If test depth alone is responsible for the above results, then the illumination in the two depth planes is irrelevant. According to the coplanar illumination hypothesis, however, the illumination difference in the two depth
planes is critical. To compare these two possibilities, we switched the luminances of selected patches that shared the same spatial location within each Mondrian so that the average luminance in the near Mondrian and the far Mondrian was approximately the same. For example, the luminances of patch A and patch A’ of Fig. 2 were switched, while the luminances of patch B and patch B’ were not. To preserve local contrast, we did not alter the patches immediately surrounding the test and the comparison. We selected the patches that were switched in order to retain the original experiment’s luminance information while minimizing the difference in average luminance between the two Mondrians. The resulting geometric mean reflectance of what had been the highly illuminated (upper and nearer) Mondrian was 13.7% (Munsell value \( v = 4.2/ \)), while the dimly illuminated (lower and farther) Mondrian was now 14.5% (Munsell value \( v = 4.4/ \)). Because each Mondrian was composed of patches of different luminances (i.e., patch luminances were ungrouped), the entire display appeared as a series of reflectances (30:1 range) under a single illuminant. As predicted by the coplanar illumination hypothesis, this manipulation eliminated the difference in brightness judgments as the perceived depth of the test was varied (Fig. 11, symbols connected by solid lines). There was no reliable effect of depth on the observers’ settings of the test-patch luminance (JS, \( p > 0.50 \); SG, \( p > 0.50 \)).

Brightness matches approximated a physical-luminance match when the luminances of the noncontiguous patches were ungrouped. This occurred despite the differences in the level of light immediately surrounding each patch (i.e., the comparison patch was five times higher). The dramatic effect of grouping the luminances of only the noncontiguous patches is demonstrated by the difference between the grouped results of Fig. 3 (replotted here as dashed lines) and measurements from this experiment (solid lines). The large effect of grouping the remote noncontiguous patches is further evidence against local edge information as the primary mechanism of brightness perception.

Experiment 10: Variation of Test Size

The test size was constant in all previous experiments so that the test appeared smaller when it was in the nearer depth plane. A 1° × 1° test in the no-depth condition is perceived to be approximately the same size as a 0.8° × 0.8° test in the depth condition. To ensure that perceived size did not affect brightness judgments, we compared measurements in the depth condition with use of a 1° × 1° test and a 0.8° × 0.8° test. There were no significant
differences in brightness judgments (JS, p > 0.35; SG, p > 0.30). Therefore a difference in apparent test size cannot account for the findings of the original experiment.

DISCUSSION

To Be Inferred, Illumination Differences Must Be Perceived

The process of judging the lightness and brightness of a surface in a three-dimensional scene requires a perceptual representation of the illuminations that fall within that scene. Since retinally noncontiguous coplanar surfaces often share a common illuminant, the coplanar illumination hypothesis makes interpretation of the luminance of any surface less ambiguous. Judging the appearance of a surface by using only local retinal luminance contrast can be misleading when retinally adjacent areas are stimulated by light from surfaces in different depth planes with different illuminations. Incorporation of information from an inferred coplanar illuminant could rectify this situation.

The coplanar illumination hypothesis requires a perceptual representation of at least two different illuminants within a scene. Selective grouping of patches with a 30:1 range of luminances made two Mondrians appear to be under different illuminations. Spatially ungrouping these patches eliminated any basis for perceiving a difference in illumination. Consequently, as predicted by the coplanar illumination hypothesis, ungrouping also eliminated a difference in the brightness judgments of the test as perceived depth of the test was varied (see Experiment 9, Mondrians with ungrouped illumination). This implies that spatial integration and segregation across the entire visual field contributes to the perceived brightness of a surface.

An important finding here is that a fivefold difference in luminances between the two Mondrians produced at most a 15% change in test luminance as the perceived depth of the test moved between the depth planes of the two Mondrians. This is far less than the 500% change in the test settings that would be predicted if the illumination of the test were inferred according to the fivefold difference in Mondrian illuminations. This discrepancy indicates that although information about the coplanar illuminant may be inferred, it is not taken as the actual illumination of the test surfaces.

Lightness

Lightness judgments also do not rely solely on luminance–edge ratios. An integration step must combine ratios from various edges to give a spatial distribution of relative lightness. Even so, proper segmentation remains a problem since reflectance and illumination edges remain confounded. Coplanarity can be a cue to whether a surface's edge ratio contains a change in illumination and reflectance.

In order for lightness constancy to be maintained, the luminance of an in-depth test should increase so that an increase in inferred illumination is offset. While the most experienced observer, JS, demonstrated this pattern with use of a gray surround, inexperienced observers SG and BP did not. Their results suggest they made local ratio matches when the test was in either depth plane. When a black surround replaced the gray one, thus eliminating local edge contrast as a cue, lightness judgments showed an expected increase in test luminance when the test was perceived in the near depth plane (compare Figs. 4 and 5). A black surround forces inexperienced observers to use a retinally nonadjacent gray scale; the result is establishment of a more-global perception of illumination differences. As with brightness judgments, however, the change in test luminance with depth was much less than what was predicted only from the coplanar illumination (observed average change of 17%, compared with 500% predicted for the coplanar illuminant).

CRT Simulation

Stereoscopic viewing of a CRT is a method of constructing three-dimensional scenes designed to simulate natural viewing of papers under two different illuminants. A CRT offers accuracy, and haploscopic viewing enhances the screen's realism by producing clearly defined depth planes. For example, stereoscopic photography has been reported to produce better lightness constancy than does a single-lens camera. Most important, the stereo displays ensured that the perceived depth information was affecting perception at a locus beyond the retina.

Comparison with Previous Studies

In 1977 Gilchrist claimed that the coplanar ratio hypothesis determined the lightness of a surface. This hypothesis tacitly assumes that the illumination that is falling across any particular depth plane is constant. Gilchrist constructed stimuli with a single inferred illuminant across at least one coplanar retinally adjacent luminance edge from which observers could derive a lightness ratio (Fig. 1 of Ref. 2). This method differs from that of the current study, in which an in-depth test patch has no retinally adjacent coplanar surround.

Lightness judgments obtained with such stimuli suggest that perceived coplanarity provides some illumination information, while noncoplanar, even nonretinally adjacent, surfaces may establish a gray scale. We can only speculate on how an observer selects an appropriate gray scale when the test's retinally adjacent borders are in a different depth plane. The lightness instructions were specifically designed to minimize the biasing of observers toward referencing the test to a particular Mondrian or illumination (see instructions to observers, above). Lightness results suggest that observers used the gray scale primarily from the retinally adjacent, noncoplanar surround. Lightness judgments were affected by the coplanar illuminant only when the retinally adjacent surround was extinguished (0% reflectance).

We agree with Gilchrist that local contrast alone is insufficient for determining lightness. However, we suggest that his observers may have inferred the illumination that was falling on the test from a retinally adjacent coplanar gray scale. This assumption underlies the hypothesis that surface lightness is determined by coplanar luminance ratios. We propose that, when the test does not share a coplanar retinally adjacent edge, the effect of a different illuminant is diminished, and noncoplanar local contrast may dominate lightness judgments.
Table 1. Ratio of the SEM for Log Brightness to the SEM for Log Lightness for In-Depth, No-Depth, and On-Average Tests for Subjects JS, SG, and BP

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Depth</th>
<th>No Depth</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gray surround</td>
<td>2.11&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.77&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.44</td>
</tr>
<tr>
<td>2. Black surround</td>
<td>2.49&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.19&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.84</td>
</tr>
<tr>
<td>4. Test and surround covary</td>
<td>2.06&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.90&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.48</td>
</tr>
<tr>
<td>5. 2D Mondrians</td>
<td>4.41&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.67&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.54</td>
</tr>
<tr>
<td>6. Test behind</td>
<td>3.48&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.91&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.70</td>
</tr>
<tr>
<td>7. Comparison behind</td>
<td>3.29&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.01&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.15</td>
</tr>
</tbody>
</table>

<sup>a</sup>p values were determined with a sign test; <sup>b</sup>p < 0.05; <sup>c</sup>p < 0.01.

**Variability**

Brightness settings tend to be significantly more variable than lightness settings. It is noteworthy that the standard error of the mean (SEM) was substantially larger for brightness measures than for lightness measures in the current study. This was the case for each observer across all experiments. Brightness variability decreased when the test was coplanar with the brightly illuminated Mondrian, except when the perceived depth of the comparison patch was varied (Experiment 8). Table 1 summarizes the brightness log SEM to lightness log SEM ratio, averaged over observers, for all relevant experiments.

The variability in lightness judgments is far less than the difference in gray-scale reflectances that might constrain the upper and lower (lighter and darker) limits of the test. The lightness SEM across experiments is approximately 1.0%, while the median difference in surface reflectance within either Mondrian was 6.0%. This suggests that constraints in addition to the upper and lower limits of any two patch reflectances restrict lightness judgments.

**Decrements**

Preliminary experiments were done with decrements in the same three-dimensional scenes in which increments were tested. Unfortunately, these findings are not easily interpreted with the framework developed here. We reversed the illumination falling on the depth planes by making the near Mondrian dimly illuminated and the far Mondrian brightly illuminated. This was done to avoid the possibility of the observer's setting the test to an increment to match a (comparison) decrement when the test was perceived in the near depth plane. Observers did not vary the luminance of the test with perceived depth when judging either brightness or lightness. Likewise, use of decrements with Mondrians under the illumination conditions of Experiment 1 also showed no effect of depth of test.

The often-noted asymmetries between increments and decrements may also apply to the Coplanar Illumination information. For example, Noguchi and Kozaki noted that the appearance of a white surface is strongly affected by changing illumination, while a black surface tends to be unchanged despite changes in illumination. Likewise, varying either the reflectance or the illumination of a test surround embedded within a two-dimensional Mondrian demonstrated a strong effect of local contrast on decrements but not on increments. Additional experiments are needed for determining whether any of these related findings explains why inferred illumination affects increments differently than it does decrements.

**ACKNOWLEDGMENTS**

This research was supported in part by National Institutes of Health grants T32 EY-07098, EY-04802, and EY-07390. We thank Sandy Guzman and Becky Powell for serving as subjects, Linda Glennie for computer programming, Cheryl Schirillo for proofreading, and Larry Clarkberg for technical assistance.

*Present address, Department of Psychology, University of San Francisco, 2130 Fulton Street, San Francisco, California 94117-1080.

**REFERENCES AND NOTES**

7. Perceived brightness and perceived lightness in a three-dimensional scene can be derived from local contrast of retinally adjacent coplanar surfaces that share an illuminant (see Ref. 8).
14. In a related paradigm, L. Kardos, "Ding und schatten," Z. Psychol. 23, (1934), claimed that surface lightness was determined by the illumination in a given depth plane. He reduced the illumination falling upon a light-gray test to make the luminance of the test approximately equal to that of a highly illuminated dark-gray coplanar surround. This caused the test both to appear dark-gray and to be under the same high illumination as its coplanar surround. Reducing the overall illumination falling upon the light-gray test and the dark-gray surround in a separate depth plane preserved the lighter test. Although Kardos used depth to segment multiple illuminations within a scene and showed that coplanar surfaces with equal luminance appear to have both equal reflectance and equal illumination, he did not derive a coplanar rule.
The two displays of Arend and Goldstein\(^{30}\) "were identical paper arrays illuminated by different sources." Their instructions to "adjust the test patch to look as if it were cut from the same piece of paper" makes explicit to the observer that the test and comparison patches were under different illuminants. The current instructions to create the "same shade of gray" do not demand that the observer infer a different level of illumination for the test and comparison. Jacobsen and Gilchrist (Ref. 3 and Ref. 16) also obtained lightness judgments by having observers match the "shade of gray.


The two displays of Arend and Goldstein\(^{30}\) showed that a small test disparity affected brightness judgments by approximately 10%, whereas addition of disparity to the test did not increase the effect.


Gibbs and Lawson\(^{32}\) observed the effect of depth separation on lightness.

Observers did not report that the test appeared in a film within the far Mondrian.


Mershon and Gogel\(^{32}\) reported a maximum perceived difference of 0.7 Munsell step when the far disk was 44% farther than the near disk. Their hypothesis of increased neural uncoupling with increasing depth implies that with sufficient separation, induction effects would disappear altogether. In the current experiment the far Mondrian was 19% farther than the test patch that was perceived in the near plane.


The fact that such a restricted range is sufficient suggests that many natural scenes potentially contain multiple levels of illumination. For example, in more than 100 natural outdoor scenes the average contrast is approximately 160:1\(^{34}\) while luminance ratios greater than 60:1 tend to be perceived as changes in illumination instead of as changes in lightness (Refs. 32 and 33 below).


R. Evans and J. Klute, "Brightness constancy in photographic reproduction," J. Opt. Soc. Am. 34, 540–553 (1944), claim that illumination differences must be apparent in black-and-white photographs for adequate lightness constancy to be obtained.


The following sources provide specific incidents of increment–decrement asymmetries: brightness and lightness,\(^{11}\) lightness and illumination,\(^{12}\) figure–ground,\(^{13}\) adaptation,\(^{14}\) qualities of "pronouncedness" (Ausgepragtheit) and "insistence" (Eindringlichkeit),\(^{15}\) illumination,\(^{16}\) and grouping.\(^{49}\)

