# Coming to terms with lightness and brightness: Effects of stimulus configuration and instructions on brightness and lightness judgments

**Barbara Blakeslee**     Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND, USA

**Daniel Reetz**     Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND, USA

**Mark E. McCourt**     Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND, USA

To recover surface reflectance and illuminance from the raw luminance signal, the visual system must use prior assumptions and strategies that make use of additional sources of information. Indeed, it has been found that depending on experimental conditions, lightness (apparent reflectance) may refer to judgments that are similar to brightness judgments (apparent luminance), that are similar to local brightness–contrast judgments, or that represent an independent third dimension of achromatic experience which exists only when the illumination across regions of the display is visibly non-uniform (L. E. Arend & B. Spehar, 1993a, 1993b). This means that *lightness* data generated in one experimental condition may not be comparable to *lightness* data measured in other conditions. We investigate this problem with regard to a history of data on simultaneous brightness-contrast by measuring brightness, brightness-contrast, and lightness in stimuli similar to those used in Gilchrist's edge-substitution studies (A. Gilchrist, S. Delman, & A. Jacobsen, 1983) and in stimuli similar to those used to test Gilchrist's intrinsic-image model against his newer anchoring model (A. Gilchrist, 2006). Our results clarify confusions that appear to stem from comparing different types of lightness judgments and from inadvertently using brightness as an index of lightness under conditions where independent lightness judgments are possible.

## Introduction

A central problem in the study of brightness/lightness perception is to understand how and under what circumstances the visual system is able to separate the physically invariant reflectances of surfaces from their potentially changing illumination. The luminance distribution falling on the photoreceptor array is the product of the reflectance of surfaces in the environment and their illumination. The recovery of surface reflectance and illuminance information from the raw luminance signal is thus an ill-posed problem in that there are myriad combinations of illumination and reflectance that can give rise to any particular intensity distribution, and in the absence of additional information there is no way to recover the correct solution. We know from everyday experience, however, that the visual system does manage to solve the inverse problem in a satisfactory manner much of the time and must, therefore, use prior assumptions and processing strategies that provide and/or make use of additional

information. Indeed, much of the current debate concerning brightness/lightness perception centers on discovering the mechanisms the visual system uses to solve the inverse problem.

One factor impeding progress in this area, however, is confusion due to definitional differences and inconsistencies in the use of the terms lightness and brightness. The CIE defines brightness as the attribute of a visual sensation according to which a given visual stimulus appears to be more or less intense, or according to which the area in which a visual stimulus is presented appears to emit more or less light (Wyszecki, 1986; Wyszecki & Stiles, 1982). Lightness, on the other hand, is defined by the CIE as the attribute of a visual sensation according to which the area occupied by the visual stimulus appears to emit more or less light in proportion to that emitted by a similarly illuminated area perceived as a "white" stimulus (Wyszecki, 1986; Wyszecki & Stiles, 1982). Although the CIE definitions of brightness and lightness are widely used, so too are definitions agreed upon by the Trieste Group (Arend, 1993) in which brightness is defined

simply as apparent luminance, and lightness as apparent reflectance. While the two slightly different definitions of brightness are not problematic, the alternate definitions of lightness have sponsored confusion. Although lightness is highly correlated with reflectance under the stimulus conditions specified by the CIE definition, in any given situation the visual system may, or may not, have access to a white stimulus under similar illumination. Lightness defined simply as apparent reflectance, however, is also problematic since it refers directly to a surface property of the stimulus (its reflectance) which, due to the inverse problem, is underdetermined. Because of this, judgments of lightness, defined as apparent reflectance, may depend on very different information and strategies under different circumstances.

The work of Arend and Spehar (1993a, 1993b) illustrates that this is indeed the case by revealing that, depending on stimulus conditions and subject instructions, lightness (defined as apparent reflectance) may refer to judgments (1) that are identical to brightness judgments (apparent luminance), (2) that are identical to local brightness–contrast judgments (apparent local luminance difference between a target and its background), or (3) that represent an independent third-dimension of achromatic experience. This result makes sense because when illumination is homogeneous, luminance and reflectance are highly correlated, making brightness judgments the best predictor of lightness. Similarly, because luminance contrast, and therefore brightness–contrast, is invariant with changing illumination level, matching brightness–contrasts offers the best strategy for matching reflectance under conditions of homogeneously changing illumination. Critically, however, Arend and Spehar (1993a, 1993b) found that lightness existed as an independent third dimension of achromatic experience only when the illumination across regions of the display was visibly non-uniform. Blakeslee and McCourt (2003) called this type of lightness judgment "inferred-lightness" because it requires that observers take account of illumination to make judgments of reflectance. In other words, unlike lightness judgments based on brightness or brightness–contrast, inferred-lightness is not a sensory-level judgment of surface color or contrast but rather requires a perceptual parsing of the sensory information into separate illumination and reflectance components, i.e., into intrinsic images (Arend, 1994; Gilchrist, 1979; Gilchrist, Delman, & Jacobsen, 1983; Kingdom, 2003). The types of judgments we are separating into "sensory-level" and "inferred" have alternatively been referred to as "phenomenal" and "projective," respectively, by other authors (Reeves, Amano, & Foster, 2008). Once this parsing of the image has occurred, judgments of lightness may be relatively easy (seemingly automatic) or difficult (clearly effortful), depending on the richness of the scene information available. For example, when a small shadow falls across a surface, thus changing its local brightness, the reflectance of that shadowed surface can be readily recovered

by identifying the lightness of the shadowed region with that of a neighboring unshadowed part. In other words, after the illumination component is identified, the task is reduced to an "easy" sensory or phenomenal judgment of the brightness of the unshadowed part, the dimension most highly correlated with lightness under these conditions. A much more "difficult" inferred or projective lightness judgment is required when the surface to be judged lies completely within (or beneath) a region of special illumination such as a shadow or spotlight, and where no similar surface outside of the special region of illumination can be readily identified with it. In this situation it is only possible to estimate (or make an educated guess about) the reflectance of the surface in question by first judging the magnitude of the illumination based on its effect on the brightnesses of remote regions—as for example where an illumination border falls across other homogeneous surfaces in the scene—and then discounting such illumination to estimate the reflectance of the surface.

The fact that lightness judgments are based on very different information under different conditions has resulted in confusion since lightness data generated in one condition or study may not be directly comparable to those measured in other conditions or studies. We investigate this issue with regard to a history of data on simultaneous brightness/lightness contrast, in which discrepancies in the magnitude and direction of the reported lightness effects appear, in some instances, to be directly attributable to comparing disparate types of lightness judgments (Gilchrist, 1979, 1988; Gilchrist et al., 1983) and, in other instances, to erroneously judging brightness rather than the available dimension of inferred or projective lightness (Gilchrist, 2006; Gilchrist et al., 1999).

In the standard simultaneous brightness/lightness contrast display, identical mid-gray targets are centered on white and black juxtaposed backgrounds (Figure 1A). Gilchrist (Gilchrist, 1979; Gilchrist et al., 1983) produced several variants of this stimulus using illumination differences, rather than reflectance differences, to produce the backgrounds. They reported that a mid-gray background was made to appear white on one side and black on the other by illuminating half of the background with a light that was 30 times more luminous than that illuminating the shadowed half (thus mimicking the intensity ratio of white to black paper under homogeneous illumination). Likewise, in order to equate the luminances of the target stimuli, as in the standard simultaneous brightness/lightness contrast display, the target on the illuminated side was black, such that it reflected 30 times less light than the white target on the shadowed side. In this way, the investigators reproduced the retinal light intensities that result from viewing a standard reflectance-based simultaneous brightness/lightness contrast stimulus under homogeneous illumination. When this illumination-edge stimulus was viewed through a rectangular aperture that

masked everything but the targets and their near backgrounds (Gilchrist, 1979)—or similarly, in a manner that obscured the actual illumination conditions (Gilchrist et al., 1983)—the display looked like a standard simultaneous brightness/lightness contrast display. The authors reported that observers described the illumination edge as a reflectance edge between black and white backgrounds and matched the lightness of the targets to mid-gray (one slightly darker than the other due to brightness induction). However, when observers viewed the display without the aperture (Gilchrist, 1979)—or when a background was added to reveal the illumination (Gilchrist et al., 1983) so that observers could clearly see and describe the differential illumination on the two halves of the stimulus—they matched the lightness of the illuminated target to black and the lightness of the shadowed target to white (Figure 1B). Thus, the lightness (apparent reflectance) of the two targets was reported to be profoundly different in this condition even though the intensity of light coming from each target and near background remained the same. Gilchrist (Gilchrist, 1979; Gilchrist et al., 1983) interpreted this result as evidence that, when given enough information, the visual system could classify edges into those due to illumination versus those due to reflectance prior to integrating the reflectance edges to determine the reflectance of various regions. Note that although only lightness matches were obtained in these experiments, subjects did report the illumination conditions that they perceived. Based on this information and on several subsequent studies (Arend & Goldstein, 1987; Arend & Spehar, 1993a, 1993b; Schirillo, Reeves, & Arend, 1990), one would expect that observers' brightness matches would have been identical to the lightness matches in the masked condition where observers reported that the illumination was homogeneous. In the condition where the illumination was visible, however, the brightness matches would have remained the same and, therefore, would have differed markedly from the reported lightness matches.

Interestingly, Gilchrist reported quite different results in more recent experiments designed to test the classified edge-integration model of lightness perception against his newer anchoring model (Gilchrist, 2006; Gilchrist et al., 1999). In these recent experiments, which we refer to as the unequal-increment experiments, a dark-gray target square was centered on one side, and a white target square was centered on the other side, of a black rectangular background. Following the application (to the half of the background containing the dark-gray target) of a clearly visible illumination component (spotlight) of sufficient intensity to make the dark-gray target within the spotlight possess the highest luminance, the dark-gray target was reported to be judged as significantly lighter than the white target outside the spotlight. Gilchrist (Gilchrist, 2006; Gilchrist et al., 1999) pointed out that this result was exactly opposite to that predicted by the intrinsic image model which posits that the lightness (apparent reflectance) of the illuminated target should be dark-gray and that the lightness of the target outside the spotlight should be white (Gilchrist, 2006; Gilchrist et al., 1999). Gilchrist interpreted this newer observation to support his anchoring model of lightness perception (Gilchrist, 2006; Gilchrist et al., 1999) as opposed to his earlier classified edge-integration model (Gilchrist, 1979, 1988; Gilchrist et al., 1983). Anchoring predicts that the lightness of the dark-gray target in the spotlight will be white because it is the highest luminance in its local framework, as well as in the global framework. The lightness of the white target outside the spotlight is predicted to be light middle-gray because, although it is the highest luminance in its local framework, it is not the highest luminance in the global framework (Gilchrist, 2006; Gilchrist et al., 1999).

Although the anchoring model has been offered as an explanation for the standard (reflectance-based) simultaneous brightness/lightness contrast effect (Gilchrist, 2006; Gilchrist et al., 1999), it has never been applied to the edge-substitution stimuli (Gilchrist, 1979; Gilchrist et al., 1983). In the standard simultaneous brightness/lightness contrast stimulus, anchoring predicts that the lightness of the target on the white background appears mid-gray because it is mid-gray relative to both the local and global
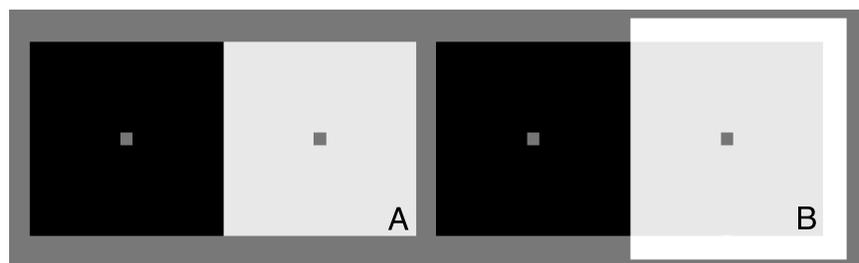


Figure 1. The two simultaneous brightness contrast stimulus configurations used in Experiment 1. The stimulus labeled A is an example of a standard simultaneous brightness contrast stimulus in which the test patch on the bright/white background has the same luminance as the test patch on the dark/black background. This stimulus resembles the masked-illumination condition of Gilchrist's early edge-substitution studies (Gilchrist, 1979, 1988; Gilchrist et al., 1983). The stimulus labeled B is identical to A with the addition of a far surround on three sides of the right-hand background. This configuration resembles the stimulus in the visible illumination condition of Gilchrist's early edge-substitution studies.

framework. The lightness of the equiluminant target on the black background, however, is predicted by anchoring to be mid-gray relative to the global framework, but white relative to the local framework, and thus to appear somewhat lighter. Using this same reasoning, anchoring can easily explain the simultaneous brightness/lightness contrast effect under the masked-illumination conditions of Gilchrist's early edge-substitution experiments, but it cannot explain the profound lightness difference reported when the non-uniform illumination is clearly visible (Gilchrist, 1979; Gilchrist et al., 1983). In this case, anchoring predicts that the lightness of the black target on the illuminated background should be darker than in the previous display because it is anchored to a higher luminance both locally and globally (the highest luminance is now the visible far surround). The lightness of the white target on the shadowed background is still predicted to be white relative to its local framework; however, the global framework now contains a higher luminance anchor than it did previously. Thus, the lightness of the test patch on the shadowed background is predicted to appear darker than in the masked condition, where it appeared slightly lighter than mid-gray, and is not predicted to appear white as reported.

Experiments 1 and 2 test the hypothesis that the profound lightness effect reported by Gilchrist (1979, 1988) and Gilchrist et al. (1983) in the early edge-substitution experiments, and the large discrepancies between the lightness judgments rendered in these edge-substitution studies and what one would expect based on results reported more recently by Gilchrist (2006) and Gilchrist et al. (1999) are due to the observers in the various experiments judging different stimulus dimensions when asked to report lightness (apparent reflectance). Specifically, we suggest that in the early edge-substitution experiments, when observers interpreted that the illumination was homogeneous, they made lightness judgments that would have been identical to brightness judgments. However, when the illumination was visibly non-uniform, the observers made inferred or projective lightness judgments. In the unequal-increment studies, we suggest that observers were again basing their lightness judgments on brightness, however, in this case they did so erroneously because the independent dimension of inferred or projective lightness was in that case actually available to them by virtue of a visible illumination component.

In Experiment 1, subjects matched the brightness (apparent intensity), brightness-contrast (apparent difference in luminance between the target and its surround), and lightness (apparent reflectance) of test patches in simultaneous brightness/lightness contrast stimuli similar to those used in Gilchrist's edge-substitution studies, but in our case the stimuli were rendered in a virtual-reality environment. Although the virtual-reality environment was intentionally simplified to produce well-controlled stimuli, concerns that such stimuli are not sufficiently realistic to provide a critical test of our hypothesis led to

Experiment 2, in which stimuli were created using calibrated Munsell papers and projected light. This allowed a closer approximation to the conditions in Gilchrist's early experiments (Gilchrist, 1979, 1988; Gilchrist et al., 1983) as well as a replication of his more recent unequal-increment experiment (Gilchrist, 2006; Gilchrist et al., 1999).

## Experiment 1

### Methods

One of the authors (BB) and two naïve observers (NP and AC) participated in Experiment 1. All subjects possessed normal or corrected-to-normal vision. Each subject provided informed consent and protocols were approved by the NDSU IRB.

Two simultaneous brightness/lightness contrast stimuli (Figure 1) were modeled and rendered to VRML using 3-D Studio Max (AutoDesk, Inc.) and were presented in virtual reality using Vizard (WorldViz, Inc.). Stimuli were presented on an NVisor SX binocular head-mounted display (NVis, Inc.) with $1280 \times 1024$ pixel resolution and a 60° (diagonal) monocular field of view. Frame refresh rate was 60 Hz and mean luminance was 40 cd/m$^2$. Pseudo-grayscale images possessing 1000 linear intensity steps were presented by employing a bit-stealing method (Tyler, Chan, Liu, McBride, & Kontsevich, 1992). Photometric calibration was performed with a spot photometer (Konica Minolta LS-110). Luminance linearity was accomplished via lookup table. The use of a virtual environment allowed us to test Gilchrist's edge-substitution stimuli (Figure 1) under highly controlled conditions. Subjects sat in a chair located in the center of the laboratory and viewed, through the head-mounted display, a virtual replica of the laboratory in which they were sitting (Figure 2). The stimulus display appeared on a virtual wall directly in front of them at an apparent distance of 362 cm. The luminance of the wall was 40 cd/m$^2$. To the right of the observers' chair was a spotlight pointed at the wall-mounted display. The spotlight appeared to be "ON" and to illuminate half of the stimulus when the small cutouts at the back of the spotlight were white as opposed to black. These cutouts were just outside the observers' field of view when looking directly at the stimulus, but slight head rotations would bring the lamp into view, along with the rest of the room environment. Subjects were instructed to glance at the spotlight to determine if it was "ON" or "OFF" before setting matches on each trial. They were also clearly instructed that when the spotlight was "ON" it was illuminating the right side of the display and that the intensity difference between the bright and dark halves of the background was due entirely to differential illumination. When the light was "OFF," however, they were
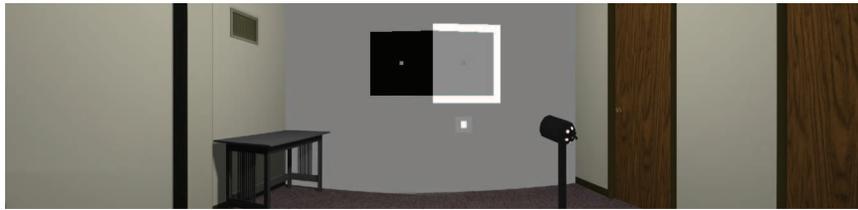
Figure 2. A panorama illustrating the virtual room (laboratory) in which the subjects viewed the stimuli of Experiment 1. Note that during the experiment, the subject was seated next to the spotlight in the virtual room and had to look down and to the right to view whether the spotlight was "ON" or "OFF."

instructed that the illumination was homogeneous and that the background intensity difference was due entirely to a reflectance change.

On each trial one of the two stimulus configurations (Figures 1A or 1B) was presented. The test patches and their near surrounds were identical in the two stimulus configurations. The square test patches subtended $1° \times 1°$ and their luminance was 40 cd/m$^2$. The near surrounds subtended $16° \times 16°$ and their luminances were 0.3 cd/m$^2$ (left-side) and 70 cd/m$^2$ (right-side). Stimulus B differed only in the addition of a far surround (80 cd/m$^2$) to the right-hand side of the stimulus. The far surround extended an additional $2°$ on three sides of the right-hand background. An adjustable matching patch ($1° \times 1°$), embedded in a square background ($2° \times 2°$), appeared below either the left or right test patch and cued the subject to match the test patch directly above it. On each trial, the background of the matching patch assumed one of seven luminance values ranging between 10 and 70 cd/m$^2$. Within a block of trials, subjects were instructed to match the brightness (apparent luminance), the brightness-contrast (apparent luminance difference between the test patch and its background), or the lightness (apparent reflectance) of the test patches. To avoid any confusion subjects were instructed that brightness matching required them to "adjust the matching patch to match the intensity or amount of light coming from the test patch ignoring, as much as possible, other areas of the display." Brightness-contrast matching instructions required them to "adjust the matching patch relative to its background to match the apparent intensity difference between the test patch and its background." Finally, lightness matching instructions required them to "adjust the matching patch to look as if it were cut from the same piece of paper as the test patch and to consider the illumination conditions in their matches." On each matching trial the initial luminance of the matching patch was randomized. Subjects controlled subsequent increments or decrements in matching luminance by rolling the wheel of a mouse upwards or downwards, respectively. Each wheel click resulted in a luminance change of 1% relative to the maximum luminance (80 cd/m$^2$). The adjustment interval for each trial lasted until the subject indicated that the match was complete by pressing a "done" button. Final adjustment settings were recorded by computer, which also randomized the presentation of stimuli. Subjects

completed 10 blocks of 112 trials in each of the three matching conditions.

## Results

Figures 3, 4, and 5 plot mean brightness, brightness–contrast, and lightness matching data, respectively. Matches for the two stimulus configurations (Figure 1) are plotted separately for each observer in the upper (stimulus A) and lower (stimulus B) panels. Within each panel, the matching patch luminance that matched the brightness, brightness–contrast, or lightness of the test patch is plotted as a function of matching background. Error bars represent 95% confidence intervals. Matches to the test patch on the dark/black background (left side) are indicated by circles; matches to the test patch on the bright/white background (right side) are indicated by squares. Filled symbols represent conditions in which the spotlight was "OFF" and subjects were instructed to interpret the brightness (apparent luminance) difference between the right and left backgrounds as due to a reflectance change. Open symbols represent conditions when the spotlight was "ON" and subjects were instructed to interpret the right half of the display as illuminated by a spotlight such that the brightness difference between the right and left backgrounds was due to an illumination change. Given these instructions there are a number of comparisons to be made across the various conditions.

First, note that matching background was varied in this experiment to obtain multiple data points and to thereby clearly differentiate between brightness, brightness-contrast, and lightness matching to the test patches of these simple stimuli. The luminance values of the left- and right-hand test patches and near surrounds remained constant throughout the experiments. The difference between the matching functions for the left- and right-hand test patches at any given matching background, therefore, represents the brightness, brightness-contrast, or lightness difference between the left- and right-hand test patch. The slope of the function generated as a function of matching background, however, reflects how the matching surround influenced the match to the test patch. Thus, as expected, the brightness difference between the left- and right-hand test patches stayed fairly constant across matching
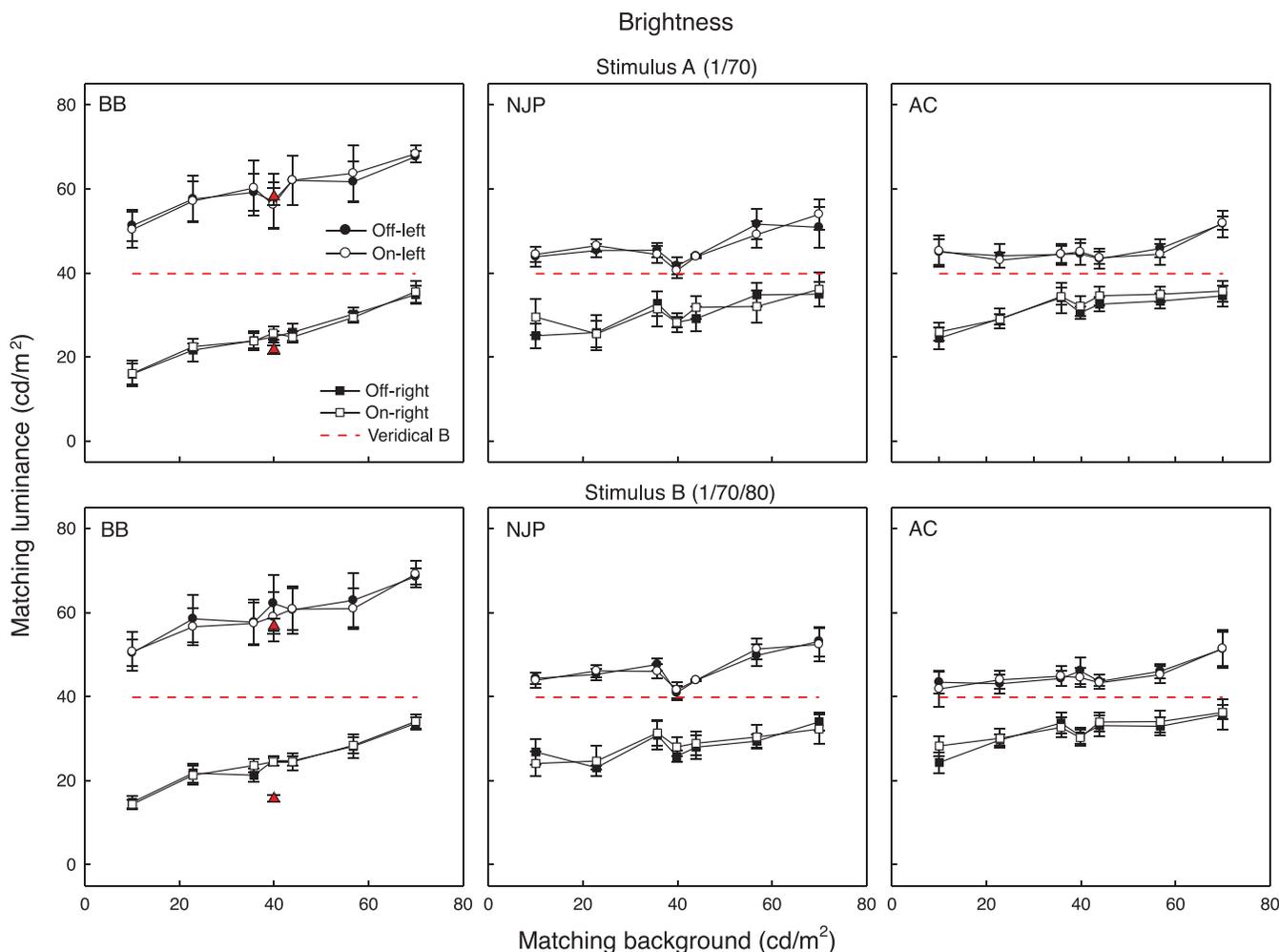
Brightness



Figure 3. Test patch brightness matching luminance (apparent luminance) as a function of matching background. Matches for the two stimulus configurations (Figure 1) are plotted separately for each observer in the upper (stimulus A) and lower (stimulus B) panels. Matches to the test patch on the dark/black background (left side) are indicated by circles; matches to the test patch on the bright/white background (right side) are indicated by squares. Filled symbols represent conditions in which the spotlight was "OFF" and subjects were instructed to assume that the brightness (apparent luminance) difference between the right and left background was due to reflectance differences. Open symbols represent conditions when the spotlight was "ON" and subjects were instructed to assume that the right half of the display was illuminated by the spotlight and that the brightness difference between the right and left backgrounds was due to illumination. Error bars are the 95% confidence intervals and the dashed red line indicates the veridical luminance of the test patches. The filled red triangles in the upper and lower left panels plot brightness matching data for observer BB obtained in the corresponding reflectance and illumination conditions of Experiment 2.

background although the magnitude of this difference varied between observers. This difference in the overall magnitude of brightness induction between observers is also reflected by the differing slopes of the matching functions. In other words, observer BB showed a large brightness induction effect between the left- and right-hand test patches and a correspondingly large induction effect of the matching background on the matching patch, as indicated by the relatively steep slope of the matching function.

Second, when the spotlight was "ON," the condition under which subjects were instructed to assume that the difference between the backgrounds on the right- and left-half of the display was due to an illumination change, subjects produced three different sets of matches for the

illuminated (right-hand) test patch (open squares) when instructed to match test patch brightness (Figure 3), brightness–contrast (Figure 4), and lightness (Figure 5). This indicates that three independent stimulus dimensions were available to be matched under these conditions. When the spotlight was "OFF," the matches to the right-hand test patch (filled squares) for both brightness and brightness–contrast measurements remain unchanged and are identical to those in the spotlight "ON" condition (open squares). The lightness (apparent reflectance) matches changed, however, and are now identical to the brightness matches. In other words, when the spotlight was "OFF," the number of stimulus dimensions was reduced and only two dimensions were available for matching. Since,
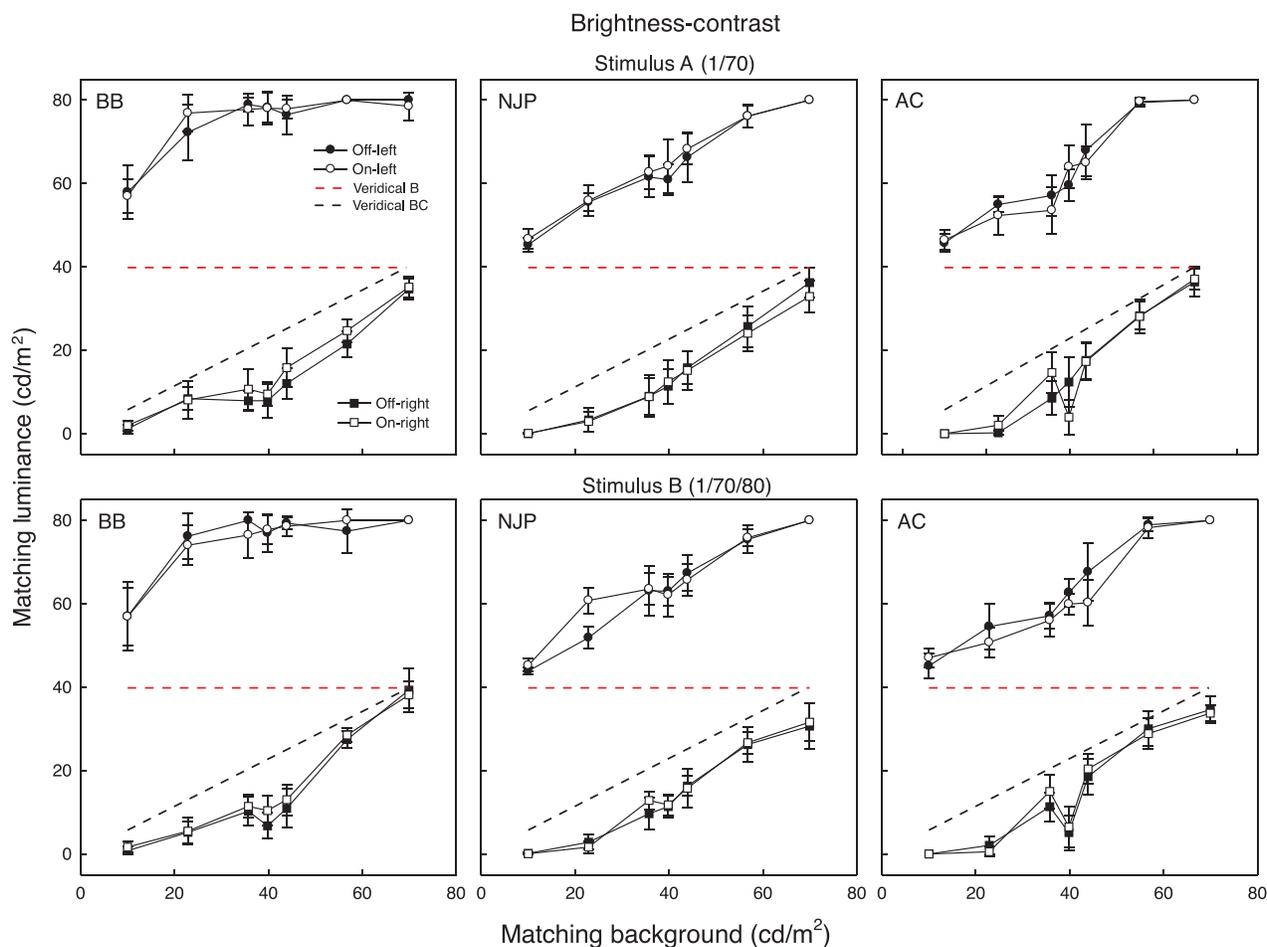
Figure 4. Test patch brightness–contrast matching luminance (apparent luminance difference between a target and its background) as a function of matching background. The black dashed line indicates the veridical luminance contrast of the test patch on the illuminated (right-hand) side. All other details are the same as in Figure 3.

under homogeneous illumination, all brightness (apparent luminance) differences are due to reflectance, it is reasonable for the two measurements to give similar results under these conditions. Importantly, however, the brightness (apparent luminance) of the test patches remained constant across the spotlight "ON" and "OFF" conditions indicating that the phenomenal appearance of the stimulus was constant, i.e., that the stimulus looked the same in both conditions, and that only the perceptual interpretation of the sensory experience was altered. In addition, because the stimuli are very simple, the inferred or projective lightness judgment was of the difficult/effortful variety discussed earlier, where the magnitude of the illumination must be estimated from the contrast information at illumination borders and then discounted from the brightness of the test patch to infer its reflectance. Subjects reported that the lightness matches when the spotlight was "ON" represented an estimate, or educated guess, based on a conscious calculation of the magnitude of the illumination component in the scene and were not sensory or phenomenal judgments of surface color. For the reasons discussed above, the lightness matches for the left-hand test patch, where illumination was identical for the

spotlight "ON" and "OFF" conditions, followed the brightness of the test patch. Finally, the similarity of the matching functions for stimulus A and B within the three matching conditions indicates that the far surround exerted very little effect on test patch brightness, brightness–contrast and lightness under these conditions. The similarity of the functions for stimulus A and B within the lightness matching condition also indicates that these stimuli were ambiguous enough that either stimulus could (with some effort) be interpreted to contain an illumination component or not, depending on the instructions given to the observer.

# Experiment 2

## Methods

The three authors (BB, DR, and MM) and three naïve observers (AC, AM, and JH) participated in Experiment 2 (AC participated only in the edge-substitution experiment
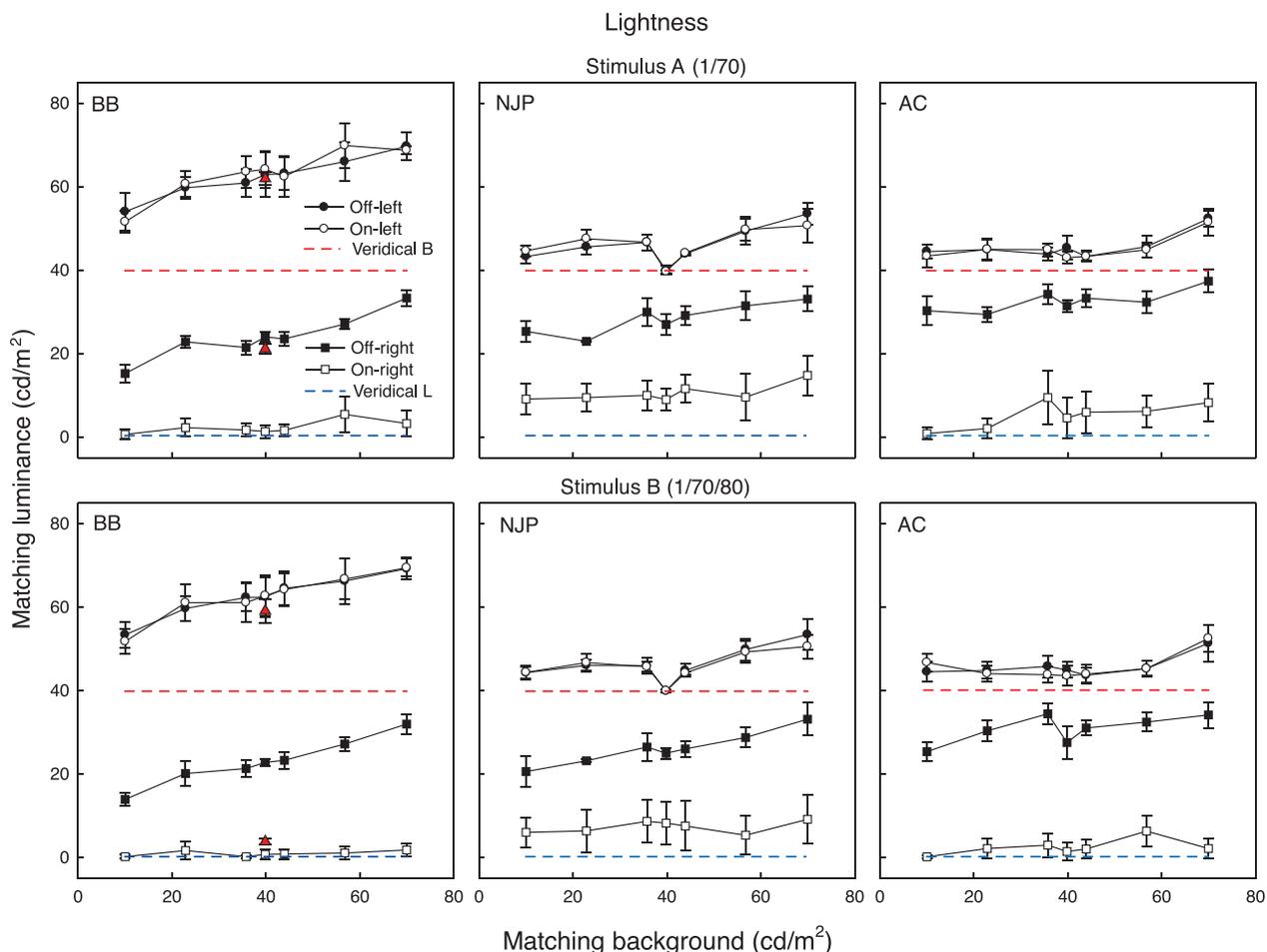
Lightness



Figure 5. Test patch lightness matching luminance (apparent reflectance) as a function of matching background. The red filled triangles in the upper and lower left panels are the lightness matching data for observer BB from the corresponding reflectance and illumination conditions in Experiment 2. The blue dashed line indicates the veridical reflectance of the test patch on the illuminated (right-hand) side. All other details are the same as in Figure 3.

and AM participated only in the unequal-increment experiment). All subjects possessed normal or corrected-to-normal vision. Each subject provided informed consent and protocols were approved by the NDSU IRB.

Stimuli were created using calibrated neutral value Munsell papers (matte) and projected light. Photometric calibration was performed using a spot photometer (Konica Minolta LS-110). Three stimulus conditions were employed to further investigate the early edge-substitution experiments (Gilchrist, 1979, 1988; Gilchrist et al., 1983) and an additional unequal-increment configuration served to replicate the more recent experimental test of the intrinsic image and anchoring models (Gilchrist, 2006; Gilchrist et al., 1999). Photographs of three of the stimuli appear in Figure 6.

Figure 6A is a standard simultaneous brightness/lightness contrast stimulus produced using Munsell papers under homogeneous illumination from a Dell 5100MP projector. An attempt was made to match this stimulus as closely as possible to stimulus A from Experiment 1. The level of homogeneous illumination was such that the

luminance of the two gray test patches was 40 cd/m$^2$, and the luminances of the black (3.1% reflectance) and white (90% reflectance) backgrounds were 3 and 80 cd/m$^2$, respectively. The luminance of the far surround bordering the backgrounds was 35 cd/m$^2$. The display was viewed from a distance of 114 cm producing test patches that measured 1° by 1°. The white and black backgrounds measured 10° × 14° and the far surround added an additional 2° around the perimeter of the display. The papers were mounted on a large (102 × 81 cm) sheet of foam core covered with black matboard. Centered at the bottom of the display was a 1° × 1° variable matching patch embedded in a 10° × 10° checkerboard background. Both the matching patch and the checkerboard were projected onto a 10° × 10° dark-gray (6.6% reflectance) surface such that the checks of the checkerboard were 20 and 60 cd/m$^2$. The matching patch was adjustable between 0.5 and 70 cd/m$^2$ in steps of 1%. The display board was attached to the wall of the lab with magnets allowing for precise registration with the projected illumination.
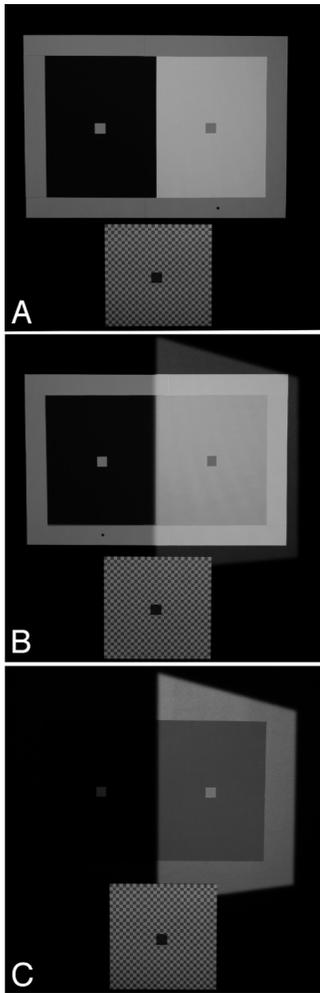
Figure 6. Photographs of three of the stimuli used in Experiment 2. (A) A standard simultaneous brightness/lightness contrast stimulus produced using Munsell papers under homogeneous illumination. (B) A simultaneous brightness/lightness contrast stimulus produced using non-homogeneous illumination to create a stimulus similar to the reflectance-based stimulus of panel A. (C) An unequal-increment display of the type used by Gilchrist (Gilchrist, 2006; Gilchrist et al., 1999) to test the intrinsic image model against the anchoring model.

Figure 6B is a simultaneous brightness/lightness contrast stimulus produced in a manner similar to that described in Gilchrist's edge-substitution experiments. In this stimulus, non-homogeneous illumination was used to create an illumination edge, which substituted for the reflectance edge of the standard stimulus. The test patches on the left and right sides of the display were cut from white (90% reflectance) and black (3.1% reflectance) Munsell papers and centered on a homogeneous gray background. The two halves of the display were differentially illuminated such that the test patches and their backgrounds possessed the same luminances as in the reflectance-based version of the display (Figure 6A). Note, however, that the illumination on the right-hand side

formed a blurred-edge trapezoidal "window" which illuminated not only the test patch, background and far surround on the right-hand side (now 90 cd/m$^2$ instead of 35 cd/m$^2$), but also some of the black matboard (11 cd/m$^2$) as well. The blurred-edge and trapezoidal-shape of the illumination were designed to serve as more salient visual cues to the illumination component of the display than were present in stimulus B of Experiment 1. A masked condition, in which subjects viewed the display in Figure 6B through a mask that blocked the illumination that fell beyond the border of the gray far surround, was also employed but is not shown. This produced a condition very similar to stimulus B in Experiment 1 and served as a control.

Figure 6C is an unequal-increment display of the type used by Gilchrist (Gilchrist, 2006; Gilchrist et al., 1999) to test his intrinsic image model against his anchoring model. In this display, a white (90% reflectance) and a dark-gray (6.6% reflectance) test patch were placed on a homogeneous black background (3.1% reflectance). On the non-illuminated (left-hand) side, the luminance of the 90% reflectance test patch was 9.5 cd/m$^2$ and the background luminance was 0.4 cd/m$^2$. The right side of this display was illuminated by a blurred-edge trapezoidal window such that the luminance of the 6.6% reflectance test patch was 49 cd/m$^2$, the luminance of the near background was 23 cd/m$^2$, and the luminance of the far background was 38 cd/m$^2$.

Because the stimuli were composed of illuminated papers and required the experimenter to change the boards and illumination for the different stimulus conditions, the experiment was run in eight blocks consisting of lightness matching or brightness matching for each of the four stimuli. The order of the blocks was such that three of the five observers performed lightness matches followed by brightness matches for each stimulus and the other two performed brightness matches followed by lightness matches. In addition, the order in which the four stimuli were presented was also counter-balanced. The observers left the room each time the stimulus conditions were changed and the illumination configuration for each new stimulus was discussed with the subject. Prior to each brightness matching block subjects were instructed to "adjust the matching patch to match the intensity of light coming from the test patch ignoring, as much as possible, other areas of the display." Prior to each lightness matching block subjects were instructed to "adjust the matching patch to look as if it were cut from the same piece of paper as the test patch and to consider the illumination conditions in their matches." On each matching trial, a small dim dot below the left- or right-hand side of the stimulus indicated which test patch was to be matched. The initial value of the matching patch was randomized. Subjects controlled subsequent increments or decrements in matching luminance by pressing buttons on a keyboard. Each button press resulted in a luminance change of 1% relative to the maximum luminance (80 cd/m$^2$). The adjustment interval for each trial lasted until the subject

indicated that the match was complete by pressing a "done" button. Final adjustment settings were recorded by computer, which also randomized which test patch was to be matched. The subjects performed five brightness matches and five lightness matches to each test patch for each of the four stimuli.

## Results

The bar graph in Figure 7 shows the means of the five observers' mean test patch matching luminances for the edge-substitution stimuli. The error bars represent 95% confidence intervals. All of the subjects showed the same overall pattern of results and are well represented by the pooled data. The dark-gray bars show the brightness and lightness matches for the left- and right-hand test patches of the standard (reflectance-edge) simultaneous brightness/lightness contrast stimulus under homogeneous illumination (Figure 6A). The white bars show the matches for the corresponding illumination-edge condition (Figure 6B) in which the test patches, their near surrounds, and the far surround on the non-illuminated (left-hand) side are identical to those in the reflectance condition. The light-gray bars show the matches in the masked-illumination condition, where the illumination falling beyond the border of the far surround was not visible.

In the standard reflectance condition (Figure 7, dark-gray bars), a simultaneous brightness/lightness contrast

effect is seen between the left- and right-hand test patches for both the lightness and brightness matches. Note that the lightness and brightness matches for the left-hand test patch are very similar as are the lightness and brightness matches for the right-hand test patch. This reflectance condition corresponds closely to the spotlight "OFF" condition of stimulus A in Experiment 1 and confirms the findings of that study. In other words, as discussed earlier, under conditions of homogeneous illumination brightness and lightness judgments are not separable (Arend & Spehar, 1993a, 1993b). The filled-red triangles plotted in the upper-left panels of Figure 3 (brightness) and Figure 5 (lightness) represent data from observer BB in Experiment 2 plotted for comparison with her data from Experiment 1. Note that the correspondence of the virtual-reality data with those obtained using Munsell papers is quite good.

In the illumination condition (Figure 7, white bars), the illumination component is clearly visible on the right-hand side of the stimulus (Figure 6B). This stimulus differs from the reflectance-based stimulus (Figure 6A) in that both the far-surround, and those areas of the matboard defining the trapezoidal shape of the illumination, are higher in luminance. As expected, the brightness and lightness matches for the non-illuminated (left-hand) side of the display do not differ from each other. However, observers produce lightness matches to the illuminated (right-hand) side which are clearly lower (darker) than the brightness matches, indicating that a separate dimension
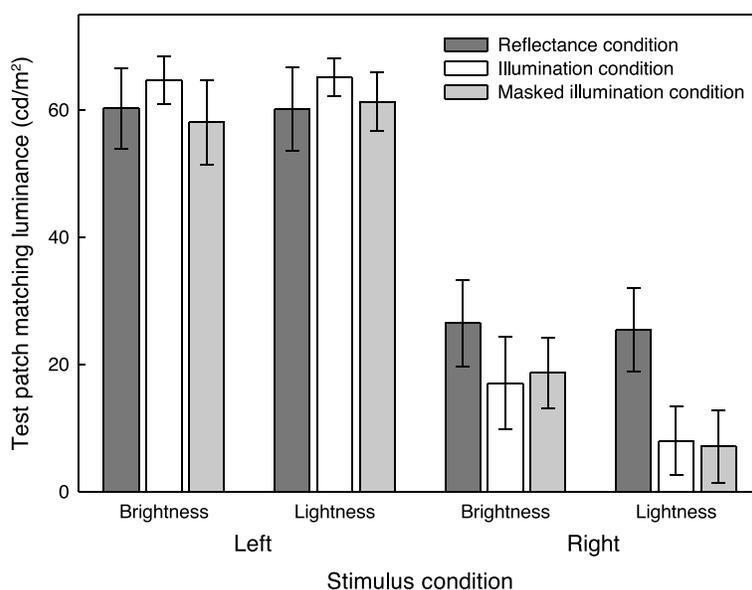


Figure 7. Bar graph showing the means of the five observers' mean test patch matching luminances, for the edge-substitution stimuli. The error bars represent the 95% confidence intervals. The dark-gray bars are the brightness and lightness matches for the left- and right-hand test patches of the standard (reflectance-edge) simultaneous brightness/lightness contrast stimulus under homogeneous illumination (Figure 6A). The white bars are the matches for the corresponding illumination-edge condition (Figure 6B) in which the test patches, their near surrounds, and the far surround on the non-illuminated (left-hand) side are identical to those in the reflectance condition. The light-gray bars are the matches for the masked-illumination condition, in which illumination falling beyond the border of the far surround is not visible.

for lightness exists under these conditions. Again, this result is similar to that obtained for the brightness (Figure 3) and lightness matches (Figure 5) to stimulus B in Experiment 1. Unlike Experiment 1, however, where no effects from the far surround in stimulus B were found, the brightness effect in the illumination condition of Experiment 2 (Figure 7, white bars) is larger than in the reflectance condition (Figure 7, dark-gray bars). This pattern of lower luminance matches for test patch brightness on the illuminated (right-hand) side was observed for all subjects. It may result from both the higher luminance (90 cd/m²) of the far surround on the illuminated side and its closer proximity to the test patch in Experiment 2. That it is not due to factors beyond the far surround is indicated by the similarity of the brightness matches for the illumination condition (Figure 7, white bars) and the masked illumination condition (Figure 7 light-gray bars). Another explanation for the exaggerated brightness difference is that brightness might be influenced in the direction of lightness (Kingdom, Blakeslee, & McCourt, 1997). However, since this effect was not observed in stimulus B of Experiment 1, this interpretation seems less likely.

In the masked-illumination condition (Figure 7, light-gray bars), the illumination-edge stimulus was viewed through a mask which blocked the spill-over of the illumination onto the black matboard, leaving the central illumination edge as the sole illumination cue. This condition was included as a control to address the concern that, in Experiment 1, the central edge by itself, even coupled with unambiguous observer instructions, might not be equivalent to a more naturalistic illumination condition. This concern appears to be unfounded, however, since again the brightness and lightness matches on the non-illuminated (left-hand) side are quite similar to

each other, and those on the illuminated (right-hand) side show an effect that is larger for lightness than it is for brightness and are of similar magnitude to the matches made in the illumination condition (Figure 7, white bars). Note that this masked-illumination condition corresponds closely to the spotlight "ON" condition for stimulus B in Experiment 1. Observer BB's data from Experiment 2 are plotted (filled-red triangles) with her data from this corresponding condition of Experiment 1 in the lower left panels of Figure 3 (brightness) and Figure 5 (lightness). Again, the data appear remarkably similar despite the somewhat lower luminance match already discussed for test patch brightness on the illuminated (right-hand) side in Experiment 2.

The bar graph in Figure 8 plots the means of the five observers' mean test patch matching luminances (brightness and lightness) for the unequal-increment stimulus (Figure 6C). The right-hand side of the black background, which contained the dark-gray test patch, was illuminated by a blurred-edge trapezoidal "window" such that the luminance of the dark-gray test patch was five times that of the white test patch on the non-illuminated (left-hand) side. As expected, the brightness and lightness matches for the white test patch on the non-illuminated (left-hand) side of the display did not differ from each other. Of more interest, however, was our finding that despite the much higher luminance of the dark-gray test patch on the illuminated (right-hand) side (49 cd/m²) compared to the white test patch on the non-illuminated side (9.5 cd/m²), its lightness was judged to be much lower. This is, of course, the expected result if the lightness judgment was based on the available independent dimension of inferred-lightness. However, this result conflicts with those reported by Gilchrist (Gilchrist, 2006; Gilchrist et al.,
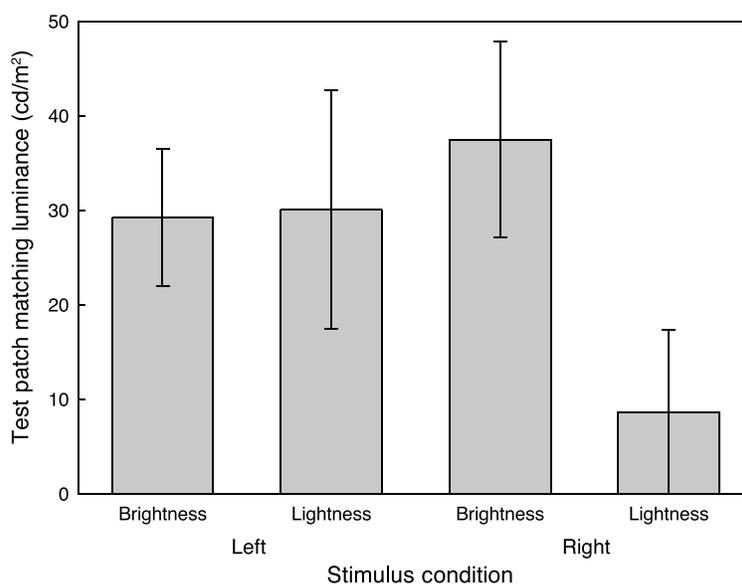


Figure 8. Bar graph depicting the means of the five observers' mean test patch matching luminances (brightness and lightness) for the unequal-increment stimulus (Figure 6C).

1999) who found that the dark-gray target within the spotlight was perceived as significantly lighter than the white target outside the spotlight.

# Discussion

The results of Experiment 1 confirm that the presence of a visible illumination component allows three dimensions of achromatic stimuli to be matched: brightness (apparent luminance), brightness-contrast (apparent local luminance difference between the target and its background), and lightness (apparent reflectance) (Arend & Spehar, 1993a, 1993b). In the absence of a visible illumination component, Arend and Spehar (1993a, 1993b) found that lightness matches collapsed onto either brightness–contrast or brightness matches depending on the specific stimulus conditions. In the present studies, the lightness matches collapsed onto the brightness matches. We conclude that when illumination is homogeneous and lightness judgments are based on brightness or brightness–contrast, they represent a sensory or phenomenal judgment of the reflectance of the stimulus that is highly correlated with these dimensions under the specific conditions of the experiment (Arend & Spehar, 1993a, 1993b). When the illumination is visibly non-uniform, however, lightness exists as an independent third-dimension of achromatic stimuli and appears to be based on an inferential or projective judgment or estimate of the reflectance of the stimulus. It is important to note that this type of lightness judgment falls outside the realm of low-level models of brightness processing and must instead be explained by higher-level inferential or perceptual models. It is only under conditions where lightness judgments are equivalent to brightness or brightness–contrast judgments that they are amenable to explanation by low-level models.

The results of Experiment 1, like those of Arend and Spehar (1993a, 1993b), indicate that care should be taken in studies of brightness/lightness perception to specify the exact stimulus conditions, to instruct subjects as to the proper interpretation of all stimuli (especially potentially ambiguous stimuli), and to provide a detailed account of all subject instructions. This practice will clarify whether *lightness* measurements in a particular study are indexing brightness, brightness-contrast, or inferred (projective) lightness and will therefore permit their comparison to lightness measurements in other studies. Furthermore, it will define which data sets should be accounted for by low-level models of perceptual processing as opposed to higher-level inferential or perceptual models.

In addition, the results of Experiments 1 and 2 support the hypothesis that the profound lightness effects reported by Gilchrist (1979, 1988) and Gilchrist et al. (1983) in their early edge-substitution experiments were likely due to two very different stimulus dimensions being matched across the two conditions. Recall that when the illumina-tion was not visible (Gilchrist, 1979, 1988; Gilchrist et al., 1983), observers reported that the illumination edge looked like a reflectance edge between black and white backgrounds. Lightness matches indicated that the targets looked mid-gray, one slightly darker than the other due to the simultaneous brightness/lightness contrast effect. This is the result seen in Experiment 1 for lightness and brightness matches (Figures 3 and 5) in the corresponding spotlight "OFF" conditions (filled squares and filled circles) for stimulus A, and for lightness and brightness matches in the reflectance condition of Experiment 2 (Figure 7, dark-gray bars)—although, in the present study the difference in the appearance of the two test patches in these stimuli would be described as substantially larger than a "slight" difference. When the illumination was clearly visible, however, Gilchrist (1979, 1988) and Gilchrist et al. (1983) reported that subjects matched the lightness of the illuminated target to black and the lightness of the shadowed target to white. In the corresponding condition (stimulus B) from Experiment 1 (Figure 5, open squares), the inferred-lightness (but not the brightness) of the illuminated target also decreased significantly, i.e., moved toward black. In Experiment 2 (Figure 7, white bars and light-gray bars) both the brightness and the inferred-lightness of the illuminated target decreased, however, the decrease was much larger for inferred-lightness. Note that, unlike in the Gilchrist experiments where the shadowed side of the display would have allowed an inferred-lightness judgment of white as reported, the non-illuminated (left-hand) side of illumination-edge stimulus in Experiments 1 and 2 did not appear to be in shadow, and therefore the brightness and lightness matches appeared the same for the non-illuminated (left-hand) side (Figure 3, stimulus B, open circles; Figure 5, stimulus B, open circles; Figure 7, white and light-gray bars). Thus, Gilchrist's reports (Gilchrist, 1979, 1988; Gilchrist et al., 1983) that the lightness (apparent reflectance) of the targets was profoundly different in the two conditions of the edge-substitution experiments, even though the target luminances were identical, seems less profound when one considers that the conditions described are those likely to produce lightness matches that are very similar or identical to brightness matches in the first instance (reflectance condition), and that represent inferred-lightness matches in the second instance (illumination condition). This point has caused a great deal of confusion in the literature as it is commonly interpreted that the lightness difference Gilchrist described was actually a huge brightness effect, i.e., an effect on a sensory or phenomenal level, which it does not appear to be (for example, see Kingdom, 2003).

The data from Experiments 1 and 2 also support the hypothesis that the large discrepancies between the lightness results in the early edge-substitution studies (Gilchrist, 1979, 1988; Gilchrist et al., 1983) and those reported more recently in the unequal-increment experi-ments (Gilchrist, 2006; Gilchrist et al., 1999) were again due to the observers in the various experiments judging

different stimulus attributes when asked to judge lightness. Recall that in Gilchrist's unequal-increment experiment, the presence of a clearly visible illumination component of sufficient intensity to make the dark-gray target on the illuminated side of the black background the highest luminance caused the dark-gray target to be perceived as significantly lighter than the white target outside the spotlight (Gilchrist, 2006; Gilchrist et al., 1999). This result was exactly opposite to that predicted by Gilchrist's classified edge-integration model (Gilchrist, 1979, 1988; Gilchrist et al., 1983) and was interpreted to support the anchoring model of lightness perception (Gilchrist, 2006; Gilchrist et al., 1999). The results from the unequal-increment stimulus in Experiment 2 (Figure 8) indicate that the brightness, but not the lightness, of the test patches followed the description in Gilchrist's unequal-increment study. In other words, the dark-gray test patch on the illuminated side had the highest luminance and appeared somewhat brighter than the white test patch on the non-illuminated side. The inferred-lightness of the dark-gray test patch in Experiment 2, however, was correctly judged to be much darker than the white test patch, despite its higher luminance. We conclude that although it was possible to judge the independent dimension of inferred-lightness on the illuminated side of the unequal-increment display, the observers in Gilchrist's studies (Gilchrist, 2006; Gilchrist et al., 1999) appear instead to have been erroneously judging brightness (the apparent luminance).

Gilchrist interpreted the results of his unequal-increment experiment as strong support for his anchoring model of lightness perception as opposed to his earlier intrinsic image model (Gilchrist, 2006; Gilchrist et al., 1999). Based on the unequal-increment results from the present study, however, it appears that a model including some analysis of intrinsic images may be required to model lightness. In addition, the present findings and recent results reported by Zdravković, Economou, and Gilchrist (2006) suggest that the anchoring model fails to predict lightness in the presence of a visible illumination component. We suggest that it might be less confusing, therefore, to consider anchoring a brightness model since, like other brightness models, it predicts lightness when lightness judgments are based on low-level sensory or phenomenal judgments of brightness (apparent luminance) or brightness contrast (apparent local luminance difference between a target and its background) but fails when there is a visible illumination component in which case lightness represents an independent dimension that must be estimated by identifying the illumination component, i.e., by parsing the image into its illumination and reflectance components, in order to determine the reflectance of the object.

## Conclusions

The present results confirm the Arend and Spehar (1993a, 1993b) finding that lightness only exists as an independent third dimension of achromatic experience when there is a visible illumination component and indicate that care must be taken in studies of lightness perception to specify the exact stimulus conditions, to instruct subjects as to the proper interpretation of all stimuli (especially potentially ambiguous stimuli), and to provide a detailed account of all subject instructions. This practice will clarify whether lightness measurements in a particular condition or study are indexing brightness, brightness-contrast, or inferred (projective) lightness and will thereby facilitate the appropriate comparisons to lightness measurements in other conditions or studies. Furthermore, it will differentiate which data sets should be accounted for by low-level sensory models of visual processing and which by higher-level inferential or perceptual models.

In addition, based on the results of Experiments 1 and 2, we conclude that the profound *lightness* effects reported by Gilchrist (1979, 1988) and Gilchrist et al. (1983) in their edge-substitution experiments were likely due to two very different stimulus dimensions being matched in the two conditions, i.e., lightness as indexed by brightness in the masked-illumination condition and by the independent dimension of inferred-lightness in the visible-illumination condition. In other words, it is clear from the results of Experiments 1 and 2 that although the inferred-lightnesses of the test patches in the visible-illumination condition were matched to black and white, they would not, as is often assumed, have appeared black and white on a sensory or phenomenal level.

We also conclude that the large discrepancies between the lightness results reported in the early edge-substitution studies (Gilchrist, 1979, 1988; Gilchrist et al., 1983) and those reported for the later unequal-increment studies (Gilchrist, 2006; Gilchrist et al., 1999) are likely the result of subjects correctly judging inferred-lightness when asked to judge lightness in the visible-illumination condition of the edge-substitution studies (Gilchrist, 1979, 1988; Gilchrist et al., 1983), but erroneously judging brightness, instead of the available independent dimension of inferred-lightness, in the later unequal-increment studies (Gilchrist, 2006; Gilchrist et al., 1999).

## Acknowledgments

Corresponding author: Barbara Blakeslee.
Email: barbara.blakeslee@ndsu.edu.
Address: Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND 58105-5075, USA.

# References

Arend, L. E. (1993). Mesopic lightness, brightness, and brightness contrast. *Perception & Psychophysics, 54,* 469–476. [PubMed]

Arend, L. E. (1994). Surface colors, illumination, and surface geometry: Intrinsic-image models of human color perception. In A. Gilchrist (Ed.), *Lightness, brightness, and transparency*. Hillsdale, NJ: Erlbaum.

Arend, L. E., & Goldstein, R. (1987). Simultaneous constancy, lightness and brightness. *Journal of the Optical Society of America A, Optics and Image Science, 4,* 2281–2285. [PubMed]

Arend, L. E., & Spehar, B. (1993a). Lightness, brightness, and brightness contrast: I. Illuminance variation. *Perception & Psychophysics, 54,* 446–456. [PubMed]

Arend, L. E., & Spehar, B. (1993b). Lightness, brightness, and brightness contrast. II. Reflectance variation. *Perception & Psychophysics, 54,* 457–468. [PubMed]

Blakeslee, B., & McCourt, M. E. (2003). A multiscale spatial filtering account of brightness phenomena. In L. R. Harris & M. Jenkin (Eds.), *Levels of perception*. Berlin: Springer-Verlag.

Gilchrist, A. L. (1979). The perception of surface blacks and whites. *Scientific American, 240,* 112–122. [PubMed]

Gilchrist, A. L. (1988). Lightness contrast and failures of constancy: A common explanation. *Perception & Psychophysics, 43,* 415–424. [PubMed]

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

Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics, 33,* 425–436. [PubMed]

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

Kingdom, F. A. (2003). Levels of brightness perception. In L. R. Harris & M. Jenkin (Eds.), *Levels of perception*. Berlin: Springer-Verlag.

Kingdom, F. A., Blakeslee, B., & McCourt, M. E. (1997). Brightness with and without perceived transparency: When does it make a difference? *Perception, 26,* 493–506. [PubMed]

Reeves, A. J., Amano, K., & Foster, D. H. (2008). Color constancy: Phenomenal or projective? *Perception & Psychophysics, 70,* 219–228. [PubMed] [Article]

Schirillo, J., Reeves, A., & Arend, L. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics, 48,* 82–90. [PubMed]

Tyler, C. W., Chan, H., Liu, L., McBride, B., & Kontsevich, L. L. (1992). Bit stealing: How to get 1786 or more gray levels from an 8-bit color monitor. *Proceedings of SPIE, 1666,* 351–364.

Wyszecki, G. (1986). Color appearance. In K. R. Boff, L. Kaufman, & J. B. Thomas (Eds.), *Handbook of perception and human performance. Sensory processes and perception* (vol. 1). New York: John Wiley and Sons.

Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae* (p. 487). New York: John Wiley & Sons.

Zdravković, S., Economou, E., & Gilchrist, A. (2006). Lightness of an object under two illumination levels. *Perception, 35,* 1185–1201. [PubMed]