Using color to understand perceived lightness

Barton L. Anderson  
Byung-Geun Khang  
Juno Kim

We report a series of experiments that employed chromatic variants of lightness illusions in an attempt to gain theoretical leverage into the types of computations that are responsible for the illusions. Two types of displays were used: one that elicits a strong percept of transparency and one that does not. We compared the pattern of induction observed in achromatic and chromatic variants of both display types. We found that the pattern of induction observed in the display that evoked a strong percept of transparency generated similar patterns of induction in both the chromatic and achromatic variants of the display, suggesting that layered image representations (scission) may be responsible for the induction observed in this display. In contradistinction, we found that the chromatic and achromatic variants of the checkerboard–gradient pattern did not consistently generate the same patterns of induction, suggesting that there are patterns of induction that only arise from luminance modulations in some class of displays. The impact of these results on existing models of lightness is discussed. We suggest that the comparison of chromatic and achromatic induction can provide insights into the computations that are responsible for the role of context in modulating perceived lightness and color.

Keywords: color vision, lightness/brightness perception, perceptual organization


Introduction

There has been a dramatic increase in the diversity of phenomena that demonstrate the effects of context on the perception of lightness and color. Yet despite the proliferation of phenomena, there is no consensus about the computations, mechanisms, or representations that are responsible for such context dependencies. One of the earliest demonstrations of the effects of context on perceived lightness and color was the phenomenon dubbed “simultaneous contrast,” which suggests both the nature of the phenomenon and its presumed cause. The idea that simultaneous contrast was the result of processes that encode contrast can be traced to Hering (1874/1964) and found modern expression in concepts of lateral inhibition (Cornsweet, 1970) and spatial filtering models (e.g., Blakeslee & McCourt, 2004). At the other end of the explanatory spectrum, Helmholtz (1866/1962) suggested that simultaneous contrast arose from processes designed to achieve color and lightness constancy. The Helmholtzian view assumes that the chromaticity and/or luminance of the surround color is partially attributed to the illumination field and thereby “discounted” from the perceived lightness or color of the targets. Helmholtz’s views have found expression, to varying degrees, in a broad range of models of lightness and color perception (e.g., Adelson, 1993; Anderson, 1997; Anderson & Winawer, 2005, 2008; Gilchrist, 1977; Kingdom, 2008; Logvinenko, 1999) but perhaps most explicitly in Barrow and Tenenbaum’s (1978) intrinsic image analysis.

A key assumption underlying intrinsic image analysis is that scene properties are recovered through processes that decompose the image into different sources (or “layers”). Each source (or intrinsic image) represents one of the physical sources of chromatic and achromatic variations in an image: reflectance, illumination, and 3D shape. This view finds some support from the (apparent) fact that these different scene attributes can be dissociated in our perceptual experience, which suggests that this information is either explicitly or implicitly represented by the visual system. Nonetheless, a number of researchers question whether layered decompositions play a causal role in the perception of lightness. For example, Gilchrist rejected his earlier layers model in favor of an anchoring model, which asserts that the highest luminance plays a privileged role in establishing absolute lightness values within and across regions that are presumed to share common illumination (Gilchrist et al., 1999). A related transformation in theoretical views is evident in Adelson’s (1993) work, which shifted from a layer model of brightness to an approach that attempts to recover lightness by identifying regions embedded in common “atmospheres,” which he defines as regions that share illumination fields and conditions of transparency. No explicit decomposition into layers occurs in either of these models.
The problem of determining whether layered image decompositions contribute to the perception of lightness inspired one of us to develop a class of stimuli that could provide direct phenomenological evidence for the role of scission if it did, in fact, play a causal role in the perception of lightness. Virtually all studies of lightness manipulate the surrounds of uniform targets and measure the influence that different contexts have on perceived lightness. In such displays, it is often perceptually difficult to determine whether the targets are decomposed into layers or not. By using textured targets, the perception of layers becomes phenomenally explicit. We demonstrated that the perception of lightness and color could, in some contexts, be modulated by the way that depth and lightness is assigned to different layers in both lightness (Anderson, 1999, 2003; Anderson & Winawer, 2005, 2008) and color displays (Anderson & Khang, 2010; Wollschläger & Anderson, 2009). These results demonstrate that scission qua transparency can induce dramatic transformations in the perception of surface lightness and color, but it remains unknown precisely when layered image decompositions are responsible for transformations in surface reflectance. The goal of the present studies was to explore whether scission plays a causal role in transforming the perceived albedo of two types of lightness induction: a display that elicits percepts of homogenous transparency and a display where the induction is caused (at least in part) by luminance gradients, for which there is no explicit evidence supporting the role of scission in the observed induction.

Our first goal was to determine whether scission could play a causal role in the lightness induction observed in displays that evoke vivid percepts of transparency that contain homogeneous targets. We have previously shown that analogous forms of induction can be observed in achromatic and achromatic textured displays that evoke inhomogeneous percepts of transparency (Anderson & Winawer, 2005, 2008; Anderson & Khang, 2010; Wollschläger & Anderson, 2009). The photometric and geometric conditions that elicit percepts of transparency are similar for chromatic and achromatic displays (D’Zmura, Colantoni, Knoblauch, & Laget, 1997; Faul & Ekroll, 2002; Wollschläger & Anderson, 2009), so it is relatively straightforward to create chromatic variants of achromatic displays that evoke percepts of transparency. If scission is causally responsible for the lightness induction observed in achromatic displays that appear transparent, then it should be possible to construct a comparable chromatic version of the display that generates analogous patterns of (chromatic) induction. Conversely, if comparable chromatic and achromatic displays generate different patterns of induction, it suggests that scission qua transparency may not be responsible for the induction observed in a lightness display.

The comparison of chromatic and achromatic displays may also shed light on the processes responsible for the induction observed in complex displays containing luminance gradients (see Figure 1). There has been considerable debate as to whether the induction observed in achromatic displays reflect processes that infer the illuminant (e.g., Logvinenko, 1999) or whether these phenomena represent some (unspecified) form of general contrast induction (e.g., Todorović, 2006). The illumination estimation hypothesis is broadly consistent with theories that invoke scission to explain transformations in perceived lightness (e.g., Anderson, 1997; Barrow & Tenenbaum, 1978; Gilchrist, 1979), at least insofar as such models invoke an explicit decomposition of the illuminant and surface reflectance. Although there is compelling evidence that the induction observed with gradients in many lightness effects are inconsistent with theories of illumination estimation that require a globally consistent light field (Todorović, 2006), such data do not rule out the possibility that luminance gradients induce local layered image decompositions in a manner analogous to those that occur in conditions of transparency. Indeed, work from our laboratory has shown that chromatic gradients can generate vivid transformations in the perceived color of a surface in conditions consistent with inhomogeneous transparency (Anderson & Khang, 2010). Chromatic induction has been observed in analogues of a variety of simple lightness displays (simultaneous contrast, White’s effect, Mach bands, assimilation, etc.), including variants of simultaneous contrast induced by chromatic gradients. Nonetheless, a number of lightness displays containing more complex arrangements of 3D shape and luminance gradients have been developed that exhibit much larger forms of induction, but there is little consensus as to the computations or mechanisms responsible for generating these effects. One example is Adelson’s checker–shadow illusion and its variants (Kingdom, 2003a; Purves & Lotto, 2003; Todorović, 2006). Some authors advocate that the induction observed in these displays suggest that the visual system either implicitly or explicitly estimates the illuminant (e.g., Logvinenko, 1999). Others have argued that such phenomena are best understood with the concepts of anchoring and frameworks (Bressan, 2006; Gilchrist et al., 1999). Yet others have argued that such phenomena reflect (at least partially) the output of relatively early spatial filtering (Blakeslee & McCourt, 1999; Dakin & Bex, 2003; Shapiro & Lu, 2011; Todorović, 2006).

Some leverage into this debate could be gained by constructing chromatic variants of the checker–shadow display. It has been suggested that the visual system is biased to interpret local variations in chromaticity as changes in reflectance, whereas variations in luminance can be interpreted as changes in material, illumination, or shape. Although chromatic changes can also occur at illumination boundaries (such as shadows), Kingdom (2008) has provided evidence that the visual system is biased to perceive chromaticity gradients as changes in
reflectance. Indeed, if a chromatic shadow generates a sufficiently strong chromatic boundary, it is often perceived as pigment (see, e.g., Kingdom, 2008). This suggests that if the context dependency of the checker–shadow illusion (and its variants) is largely driven by processes involved in illumination estimation, then it should be difficult to generate analogous chromatic variants of the checker–shadow display that exhibit a similar pattern of induction. However, if local contrast enhancement or filtering processes are predominantly responsible for the achromatic checker–shadow illusion, then it should be possible to generate chromatic variants of these displays that

Figure 1. Checkerboard patterns used in Experiment 1. (A) Checkerboard–gradient, (B) checkerboard-only, and (C) gradient-only patterns. Each row depicts the 5 modulations: achromatic, red–green, half-contrast red–green, red-only, and green-only. The contrasts of the first-column patterns are two times larger than those of the second- and third-column patterns. Observers’ settings are shown in the fourth column (D) in which the height of bars represents the magnitude of the effects, while the color of bars indicates the color of checks surrounding the test. Differences between the two surround conditions are shown in fifth column (E).
exhibit comparable patterns of induction. A similar prediction can be made for scission models that rely on local variations in contrast along contours to decompose images (e.g., Anderson, 1997; Anderson & Khang, 2010; Anderson & Winawer, 2005, 2008). Such models would also predict that it should be possible to generate equivalent induction in chromatic and achromatic displays, since the decomposition is driven largely by local variations in contrast along contours and textures.

To gain insight into these issues, we chose to evaluate two types of achromatic displays that have been shown to exhibit strong induction effects: one that evokes a vivid percept of transparency and one that does not. For the former, we chose Adelson’s “tips and strips” display; for the latter, we chose Todorović’s (2006) checkerboard–gradient display (a variant of Adelson’s checker–shadow illusion). Both of these patterns can be conceptualized as a superposition of multiple contrast components, each of which may contribute to the induction observed in these displays (see Figures 1–3). This allowed us to measure the contribution of each component of contrast modulation and configuration separately and compare the overall pattern of results for both chromatic and achromatic variants of the displays. The logic of these experiments was that if similar mechanisms underlie the chromatic and achromatic variants of these displays, then the effect sizes should be ordered similarly across the different display types. We chose the tips–strips pattern because it evokes a vivid percept of (homogeneous) transparency throughout the targets, and because the large transformation in the tips–strips patterns seems difficult (or impossible) to reconcile with any form of contrast induction. Although it has been shown that scission qua transparency can induce dramatic transformations in perceived lightness and color in displays that evoke percepts of inhomogeneous transparency, there is little evidence that this type of scission induces strong lightness or color transformations in homogeneous targets (Kingdom, Blakeslee, & McCourt, 1997; Ekroll, Faúl, Niedereé, & Richter, 2002). The tips–strips pattern seems like an excellent candidate to find such effects, since it does not seem possible to understand these phenomena on the basis of low-level induction processes, or through other models of lightness perception (see General discussion section below).

We chose Todorović’s (2006) checkerboard–gradient variant of Adelson’s checker–shadow illusion to assess the role of scission in displays containing gradients. We reasoned that if the achromatic induction was the result of low-level forms of contrast enhancement, then the chromatic and achromatic stimuli should exhibit the same ordering of effect sizes. A similar prediction could be made for local scission accounts that rely on the contrast or intensity variations along contours and textures to decompose the image. If similar effects are observed with achromatic and chromatic variants of the checkerboard–gradient display, it would limit the explanatory scope of illumination estimation models for displays of this type (under the assumption that chromatic variations are biased to be interpreted as material changes).

**Experiment 1**

**Methods**

**Observers**

Eight observers participated in the experiment, all of whom had normal or corrected-to-normal visual acuity and normal color vision. Seven of the observers were naive to the purpose of the experiment, and one was author BK.

**Stimuli**

The stimulus used in Experiment 1 was a checkerboard–gradient pattern (Kingdom, 2003b; Todorović, 2006), which is a variant of Adelson’s checker–shadow illusion. The checkerboard–gradient pattern can be constructed by superimposing a 2D checkerboard with a spatially periodic gradient. The checkerboard-only pattern is a display introduced by De Valois and De Valois (1988). To measure the induction of each contrast component separately, three patterns were constructed: a checkerboard–gradient pattern (Figure 1A), a checkerboard-only pattern (Figure 1B), and a gradient-only pattern (Figure 1C). The checkerboard-only pattern was \( 6^\circ \times 5^\circ \), composed of alternating checks subtending \( 1^\circ \times 1^\circ \) width and height. The gradient-only pattern was a one-dimensional, horizontal trapezoidal wave of 0.25 cpd with equal length of plateau and linear ramp. The checkerboard–gradient pattern was constructed by superimposing the checkerboard and gradient patterns in such a way that the height of the checks was matched to that of the plateau or ramp of the trapezoidal wave. Test checks were regions of increment or decrement checks (relative to their surrounding checks) in the second or the fourth row of a given pattern, and the matching check was the same size of a test check, presented on the right-hand side of the display on a random-dot background. The entire display subtended \( 18^\circ \times 12^\circ \) of visual angle. The background was filled with a black and white random-dot pattern, with each dot subtending \( 1' \times 1' \). The displays were viewed from a distance of 96 cm and the physical size of the stimulus was \( 9.0 \times 7.6 \) cm, which a check size of \( 1.8 \times 1.8 \) cm.

Five different intensity modulations were treated for each of the three patterns: achromatic (luminance-only), full red–green, half-red/half-green, red-only, and green-only. The red- and green-only modulations were included to determine whether any differences that arose in the achromatic and chromatic effects could be attributed to the observation that a target region cannot appear...
simultaneously red and green, whereas achromatic transparent regions can appear as a mixture of “dark” and “light.” The half-red/half-green modulation was added to ensure that any differences between the red-only and green-only, and the full-contrast pattern, could not be attributed to the lower contrast of the former patterns.

For all five modulations, the contrasts of the two component patterns—checkerboard-only and gradient-only—were set to half of the composite checkerboard–gradient pattern. Luminance in achromatic modulation ranged from a maximum of 55 cd/m$^2$ to a minimum of 5 cd/m$^2$ with the mid-gray (origin) at 30 cd/m$^2$. For the equiluminant chromatic modulations, all chromaticity values were taken from the MacLeod and Boynton (1979) equiluminant color diagram that is a transform of the Smith and Pokorny (1975) cone fundamentals. In this space, the chromaticity of a color is described by a vector defined by two components that specify the relative cone excitations: the long-wavelength-sensitive cone excitations, $L/(L + M)$, and the short-wavelength-sensitive cone excitations, $S/(L + M)$. The CIE daylight D (D65) was taken as the origin of this space, whose chromaticity was (0.6551, 0.0169) at a luminance of 30.0 cd/m$^2$. In this color space, the saturation of a color is represented as the distance from the origin (gray), whereas the hue is described as the angular distance from a reference axis (the positive $L/(L + M)$ in the present study).

Highly saturated red and violet on the two axes, chosen suitably considering the gamut of the display, were used for references to compute chromaticity distance and contrast, whose chromaticities were (0.7000, 0.0169) and (0.6551, 0.0314), respectively. All chromaticities used in the present study were normalized by these two colors: first, subtracted by the chromaticity of the origin (D65), and then divided by two difference components, respectively, between the reference and the origin on each axis. To ensure that there were no substantial differences in equiluminance settings between individuals, we used the minimum flicker and minimum motion techniques to obtain equiluminance settings of the two highly saturated red and green colors for each participant (Wagner & Boynton, 1972). Although deviations from the CIE standard observer’s setting (30 cd/m$^2$ in the present study) were thought to be inconsequential for all participants (less than 1 cd/m$^2$), idiosyncratic luminance values were corrected per participant when stimuli were presented. In case of achromatic modulation, a luminance value of 25 cd/m$^2$ was used as a denominator for normalization. The three phosphor guns of the CRT were calibrated and had $2^{12}$ values for each pixel.

**Procedure**

The test pattern varied over the type of pattern, the type of modulation, and the polarity (whether a test was increment or decrement depending on surrounding intensity). On each trial, a test pattern was chosen randomly and presented on the center of the left half of the display, a matching check on the center of the right half of the display, and the rest of the display was filled with achromatic random dots whose mean was mid-gray. A method of adjustment was used. The observer was required to adjust the perceived color or lightness of a matching check to that of a test check within the pattern, as specified by a missing square on the row to be assessed. Stimuli remained on the screen until the observer reached a satisfying adjustment. All observers did two blocks of 30 test trials combining 3 pattern types, 5 intensity modulations, 2 surrounds (whether a given test was surrounded by higher or lower contrast surrounds). Five sessions were run for each observer, which each took approximately 45 min.

**Results and discussion**

The results of Experiment 1 are presented in Figures 1D and 1E. In Figure 1D, observers’ matches were averaged separately for each of the three pattern types, five intensity modulations, and the increment vs. decrement conditions. Figure 1E shows the pattern of difference scores for the two target types (increments versus decrements). In the insets of the two graphs, the horizontal axis represents three different pattern types and test polarity relative to the surrounding checks (increment vs. decrement). The bar color (Figure 1D) indicates the color of the surround, and the vertical axis indicates the match contrast for the corresponding display shown in Figure 1A. For the purpose of the discussion that follows, we will focus on the pattern of difference scores depicted in Figure 1E, which capture the difference in perceived lightness or color of the two test checks within the display as measured by our matching task and thus represents the size of the induction observed in each display.

The results of achromatic checkerboard pattern reveal that the effect was largest for the full checkerboard–gradient pattern, with the gradient-only pattern eliciting about half of the induction effect observed with this pattern. For the achromatic stimuli, a one-way repeated-measures ANOVA performed on the magnitude of differences in lightness matches was found to be significantly different across stimulus conditions ($F_{2.14} = 59.87, p < 0.0000005$). A series of follow-up Bonferroni-corrected paired comparisons showed that differences in lightness matches between the two checks were larger for the checkerboard–gradient condition compared with the checkerboard-only condition ($p < 0.05$). Lightness matches made in the checkerboard–gradient condition were also more biased in comparison to those in the gradient-only condition ($p < 0.05$). The differences between lightness matches made in the gradient-only condition were also larger in comparison to those made
in checkerboard-only condition \((p < 0.05)\). The results obtained with the achromatic checkerboard pattern reveal that the effect was largest for the full checkerboard–gradient pattern, with the gradient-only pattern eliciting about half of this effect.

The chromatic variants of these displays did not consistently replicate the pattern of results obtained with the achromatic stimuli. For the full-contrast chromatic stimuli, a one-way repeated-measures ANOVA performed on the magnitude of differences in color matches was found to be significant across stimulus conditions \((F_{2,14} = 7.82, p < 0.01)\). Bonferroni-corrected paired comparisons showed that differences in color matches across changes to surround intensity were not significantly different between the checkerboard–gradient and checkerboard-only conditions \((p > 0.05)\). There was also no statistical difference between the biases in matches obtained in the checkerboard–gradient condition compared with the gradient-only condition \((p > 0.05)\). However, the overall differences in color matches obtained in the gradient-only condition were significantly larger than those in checkerboard-only condition \((p < 0.05)\).

For the half-contrast chromatic stimuli, a one-way repeated-measures ANOVA performed on the magnitude differences was found to be significant across stimulus conditions \((F_{2,14} = 19.91, p < 0.0001)\). Bonferroni-corrected paired comparisons showed that differences in observer matches across changes in surround intensity were larger for the checkerboard–gradient condition compared with the checkerboard-only condition \((p < 0.05)\). Differences between matches in the checkerboard–gradient condition were also larger than those in the gradient-only condition \((p < 0.05)\). The differences between matches in the gradient-only condition were also larger in comparison to those in checkerboard-only condition \((p < 0.05)\). These data obtained with the half-contrast chromatic stimuli most closely resemble the effects observed with achromatic stimuli.

The red-only and green-only conditions exhibited some differences to the data obtained with the achromatic stimuli and half-contrast chromatic stimuli. A one-way repeated-measures ANOVA performed on the magnitude differences in the red-only condition was found to be significant across stimulus conditions \((F_{2,14} = 8.26, p < 0.005)\). Bonferroni-corrected paired comparisons showed that differences in color matches between checks in the checkerboard–gradient condition were not significantly different from matches in the checkerboard-only condition \((p > 0.05)\). There was also no statistical difference between the differences in matches obtained in the checkerboard–gradient condition compared with the gradient-only condition \((p > 0.05)\). However, the overall differences in color matches obtained in the gradient-only condition were significantly larger in comparison to those in the checkerboard-only condition \((p < 0.05)\). A one-way repeated-measures ANOVA performed on the magnitude differences in color matches made in the green-only condition was also found to significantly vary across stimulus conditions \((F_{2,14} = 34.07, p < 0.000005)\). Bonferroni-corrected paired comparisons showed that differences in observer matches across changes to surround intensity were larger for the checkerboard–gradient condition compared with the checkerboard-only condition \((p < 0.05)\). However, there was also no statistical difference between the biases in color matches obtained in the checkerboard–gradient condition compared with the gradient-only condition \((p > 0.05)\). The overall differences in color matches obtained in the gradient-only condition were significantly larger in comparison to those in the checkerboard-only condition \((p < 0.05)\).

In addition to comparing the differences across conditions, we conducted further single-sample \(t\)-tests to determine whether the difference between perceptual matches made in the checkerboard-only condition were significantly different to zero. These analyses were found to be significant for all conditions, including achromatic stimuli \((t_7 = 6.39, p < 0.0004)\), full-contrast chromatic stimuli \((t_7 = 5.61, p < 0.0008)\), half-contrast chromatic stimuli \((t_7 = 4.03, p < 0.005)\), red-only stimuli \((t_7 = 4.49, p < 0.003)\), and green-only stimuli \((t_7 = 3.33, p < 0.013)\). All of the effects were consistent with the target being driven in a direction of contrast from its surrounding checks. All of the checkerboard–gradient and gradient-only patterns were also significantly different than zero.

The differences in perceived lightness/color were always greater for the gradient-only condition compared with the checkerboard-only condition for all of the conditions. Differences in perceived color/lightness were greatest in the context of the checkerboard–gradient stimuli compared with the gradient-only stimuli in the achromatic condition and the half-contrast chromatic condition. However, this superiority of the checkerboard–gradient condition to influence perceived lightness was not observed in the full-contrast chromatic condition and the conditions with red-only and green-only stimuli. In case of red-only and green-only modulations, the effects are larger when the mid-red or mid-green target was embedded in the fully saturated surrounds (Figure 1D). This might be due, in part, to the fact that the overall mean of the test pattern was at the mid-red or mid-green, whereas the matching disk was surrounded by mid-gray and, hence, were subjected to different patterns of local adaptation.

The results of this experiment suggest that the achromatic and chromatic variants of the checkerboard–gradient pattern do not exhibit the same pattern of effects. In the full-contrast chromatic pattern, the effect of the gradient alone was actually larger than that of the checkerboard–gradient pattern, which is an order reversal relative to the order of effect sizes in the chromatic pattern. For the red and green displays, there was no significant difference between the gradient-only display and the checkerboard–gradient display. The only pattern that exhibited the same general
pattern of difference scores was the half-chromaticity display. Thus, taken as a whole, the pattern of results suggest that chromatic displays do not, in general, exhibit the same pattern of induction as observed in the achromatic displays. However, to this point, we have only assessed induction effects that occur along an opponent axis in color space. In order to assess the generality of this result, we conducted a second experiment that evaluated the role of context on color appearance for displays that varied in hue.

## Experiment 2

### Methods

#### Observers

Four observers participated in the experiment. Three of four were new observers, and one was author BK.

#### Stimulus and procedure

In this experiment, we constructed hue variants of the three patterns used in Experiment 1. The hue-modulated checkerboard–gradient patterns varied in 8 hue pairs of highly saturated colors taken along the perimeter of the hue “circle” (separated 90° in hue angle). Intermediate colors were selected so that they fell along a straight line in MacLeod–Boynton color space between the two extreme hue values (e.g., red–violet in the top row of Figure 3). Eight hue pairs that differed 45° in hue angle were tested, such that the whole hue space was evaluated. The color of the test was the mean of a given hue pair. As in Experiment 1, the square used for the matching pattern was placed on a random-dot background with a neutral (gray) mean chromaticity. The same three pattern types were used as in Experiment 1 (checkerboard–gradient, checkerboard-only, and gradient-only). Thus, there were 3 pattern types, 8 pairs of hue modulations, and 2 increment and decrement test conditions. All aspects of the procedure were identical to those of the checkerboard experiment (Experiment 1). The matching procedure was the same as that in Experiment 1, except that observers now have access to the full hue–saturation circle in selecting their match colors with the mouse.

### Results and discussion

The results of Experiment 2 are shown in Figure 2. The left side of the figure depicts the chromaticities and pattern type for each of the data sets on the right, which are plotted in the MacLeod and Boynton (1979) equiluminant color diagram. The longer colored line depict the hues used in each of the three pattern types, and the central dot located on the line represents the color of the target patch. The two dots connected to the central dot represent observers’ average matches to each display. The first noticeable feature of the data is that observers’ settings do not fall on the chromaticity line used to construct the displays but, rather, are all displaced toward less saturated colors. These settings may arise from the asymmetry in the chromatic context of the test and match, as the match was placed on an achromatic random-dot surround, and the targets were embedded in highly saturated colors of varying hues. This would cause the same hue to appear more saturated on the neutral background, presumably due to differences in local states of chromatic adaptation. This was verified with a control experiment where the random-dot surround was replaced with the extreme hue values used in the test. In this condition, observers’ matches fell along the hue lines used in the test. The second and most relevant feature of the data is that all three patterns generated substantial induction effects. Importantly, however, there was no consistent difference between the checker-only pattern, gradient-only, and the checker–gradient matches. Thus, the pattern of matches for chromatic versions of the checker–gradient images is consistently different than the pattern of matches for the analogous achromatic versions of these three display types, consistent with what we observed in Experiment 1.

## Experiment 3

The results of the first two experiments reveal that the pattern of induction observed with chromatic and achromatic variants of the checker–gradient patterns are quite different. These differences may reflect the different role of chromatic and achromatic boundaries and gradients in natural scenes. Chromatic boundaries and gradients almost always arise from variations in surface reflectance, whereas luminance edges can be caused by changes in surface orientation, albedo, or illumination. Although chromatic edges can arise from illumination boundaries (due to chromatic differences in secondary and primary illumination sources), these differences are typically relatively modest. One domain where we might expect to observe similarities in chromatic and achromatic processing is in conditions of transparency. Transparency perception involves decomposing an image into explicitly perceived surface layers, which can vary in either chromaticity or albedo. If this decomposition is responsible for the induction observed in an achromatic display, it is reasonable to expect that its chromatic analogue should exhibit a similar pattern of induction (Anderson, 1997; Anderson & Winawer, 2005, 2008; Wollschlager & Anderson, 2009). This suggests that it may be possible to determine whether scission qua transparency plays a causal role in lightness displays by constructing chromatic variants of the same display. Since it has been shown that equiluminant displays can evoke percepts of transparency,
Figure 2. Hue modulation patterns. A pair of two chromaticities on the MacLeod and Boynton (1979) equiluminant color diagram was used for the modulation of chromaticity, which was varied 90° in hue angle (e.g., red–violet) for checkerboard–gradient (the first column) and 45° for checkerboard-only and gradient-only (the second and third columns). Eight pairs of chromaticities, differing in hue angles by 45° (8 rows above), were tested. The maximum chromaticities were 0.7000 on red–green (L/(L + M)) axis and 0.0250 on violet–yellow (S/(L + M)) axis with respect to the origin of D65 (0.6551, 0.0169, CIE daylight D). The first-column patterns had two times larger contrast than the second- and third-column patterns. Observers’ settings are shown in the fourth, fifth, and sixth columns, correspondingly. In each inset, the middle point is of the chromaticity of the test, and the two horn-like points are observers’ adjustments.
we constructed chromatic variants of a robust lightness illusion that evokes strong percepts of transparency (Adelson’s tips and strips illusion) to test this hypothesis.

**Methods**

**Observers**

Fifteen observers, including the same eight observers that participated in Experiment 1, participated in Experiment 3, all of whom had normal or corrected-to-normal visual acuity and normal color vision. Four of the observers were naive to the purpose of the experiment, and one was author BK.

**Stimuli and procedure**

The images used for the display were the “tips and strips” pattern created by Adelson (unpublished; see Figure 3). We created three test displays to assess the pattern of induction caused by the different components of the display: a diamond figure with dark or light tips on a homogeneous background (Figure 3A); the full tips and strips display, consisting of the same diamonds on a horizontally striped dark and light background (Figure 3B); and homogeneous diamonds on a striped background (Figure 3C). The three patterns differ in the extent to which they give rise to a transparency percept, with the tips–strips figure generating a strong percept of transparency (Figure 3B), the tips–no strips creating a weak or absent percept of transparency (Figure 3A), and the homogeneous (no tips) pattern generating no percept of transparency (Figure 3C). For the three types of patterns, the luminance and colors were modulated in the same way as the checkerboard patterns used in Experiment 1: one achromatic display and four equiluminant color variations. The test image was the diamonds on the second or third strips, which was indicated by a marker that appeared on each trial. Observers adjusted a match pattern on an achromatic random-dot background until it appeared the same as the specified target. There were 3 pattern types, 5 intensity modulations and 2 increment or decrement test conditions. The entire pattern was 7.6 × 9.0 cm, the diamonds were 1.8 cm × 1.8 cm, and the strips were 7.6 cm × 2.2 cm, all of which were viewed from a distance of 96 cm. All other aspects of the procedure were identical to those of the checkerboard experiment.

**Results and discussion**

Observers’ matches for the tips and strips patterns are summarized in the same way as those of the checkerboard patterns and drawn in Figure 3D. The matching difference scores, which depict the overall difference in appearance of the two target types, are shown in the bar plots of Figure 3E. For all 5 modulations—achromatic, full red–green, half-red/half-green, red-only, and green-only—the tips and strips display produced a larger effect than the other two patterns (the tips–no strips and no tip–strips). One-way repeated-measures ANOVAs performed on the magnitude of differences between lightness/color matches with different surrounds was found to be significant across the three stimulus conditions in the achromatic condition ($F_{2,28} = 9.17, p < 0.001$), the full-contrast chromatic condition ($F_{2,28} = 9.63, p < 0.001$), the half-contrast chromatic condition ($F_{2,28} = 9.44, p < 0.001$), the red-only condition ($F_{2,28} = 9.35, p < 0.001$), and the green-only condition ($F_{2,28} = 7.24, p < 0.005$). A series of Bonferroni-corrected paired comparisons conducted within each condition consistently found that differences in perceived lightness/color matches to patches in the tips and strips condition were significantly larger in comparison to matches made to stimuli with tips alone ($p < 0.05$) or strips alone ($p < 0.05$). The post-hoc comparisons also consistently showed that there was no significant difference in the mean magnitudes of biases in observer matches between the tips-only and strips-only conditions ($p > 0.05$). The pattern of difference scores reveals that the same general pattern of color/lightness induction was observed across both the chromatic and achromatic patterns: The strongest induction was observed for the pattern that evoked percepts of transparency and was weakest for the patterns that did not. This suggests that similar processes are involved in the induction observed with both the chromatic and achromatic display types.

**General discussion**

The goal of the preceding experiments was to use color to gain theoretical leverage into the processes responsible for the effects of context observed in two types of lightness displays. The primary motivation for these experiments was to determine if and when scission plays a role in the effect of context in modulating perceived lightness. Our previous work has shown that similar forms of induction are observed with chromatic and achromatic patterns when the images support percepts of transparency (Anderson, 1997; Anderson & Khang, 2010; Wollschläger & Anderson, 2009). These empirical results are also bolstered by general theoretical arguments, which reveal that there are strong generative and perceptual links between the conditions that elicit chromatic and achromatic transparency (D’Zmura et al., 1997; Faul & Ekroll, 2002). The most compelling displays that demonstrate the role of transparency on lightness have been in textured displays that either evoke percepts of inhomogeneous transparency or an inhomogeneous transparent layer (Anderson & Khang, 2010; Wollschläger & Anderson, 2009). It is unknown whether scission plays a role in a variety of strong lightness illusions that also evoke percepts of transparency. The
Figure 3. Tips and strips pattern used in Experiment 3. (A) The tips display, (B) the tips and strips and strips display, and (C) the strips display. Each column depicts 5 intensity (luminance or saturation) modulations: achromatic, red–green, half-red/half-green, red-only, and green-only in the order of row. For each pattern, the second diamond on the second or third strip was used as the test. Observers’ settings are shown in the fourth column. Yellow bars in the fifth column indicate differences in the inductions between when the test was on the second row and when it was on the third row.
issue concerns whether the transparency in these displays plays a causal role in generating the transformation in perceived lightness or whether it is merely epiphenomenal. We reasoned that a minimal condition that must be satisfied for a scission theory to account for a lightness phenomenon is that it should be possible to generate an equivalent color version of that display, at least for the kind of scission that accompanies percepts of transparency. The results of the tips and strips experiment reported here provides evidence that similar effects can be observed in lightness displays using homogeneous targets and is consistent with a view that layered image decompositions contribute to the contextual effects on lightness and color that are observed in this display.

One of the reasons that we chose the tips–strips pattern was that it is highly unlikely that filtering or contrast models could explain the effects in these displays. For example, the high-pass filter model of Shapiro and Lu (2011), which is the easiest to test, predicts essentially no difference between the strips and the tips–strips pattern, which is clearly not consistent with either our data or phenomenology. This is true of virtually any simple contrast model; the contribution of the “tips” to the local contrast of the diamonds is largely negligible. In fact, the tips that border the diamonds should actually drive the illusion in the opposite direction than is experienced: The diamonds perceived as lighter should be made to appear darker and vice versa.

The results of the tips–strips experiment are also difficult to reconcile with anchoring theories. Although it is unclear how such theories can be generalized to color, in their current form, the model proposed by Gilchrist et al. (1999) would predict a significant asymmetry in the induction observed in these displays for the targets that are local increments and those that are decrements. Our achromatic data exhibit a nearly perfectly symmetric pattern of induction, which is difficult to reconcile with an anchoring model.

Although we have focused on the fact that the order of the effect size in the tips–strips pattern is the same over all of the achromatic and chromatic conditions, the strips pattern appears to exhibit proportionally more induction in the chromatic conditions than in the achromatic display. However, this difference only achieved statistical significance in the green-only pattern. To the extent that this trend represents something statistically reproducible, it is most likely due to the fact that the small T-junctions between the diamonds and the strips were much less visible in the chromatic display than in the achromatic display. This difference is due, at least in part, to the small chromatic contrasts that are achievable on our monitors.

The results of the tips–strips experiment suggest that scission is responsible for modulating the perceived lightness and color of targets in displays that evoke percepts of transparency, but the role of scission in perceived lightness in non-transparency displays remains a topic of debate. Whereas percepts of transparency are phenomenally layered, the same is not (or need not be) true for the computations that separate the contributions of surface reflectance from the illumination field. Anchoring theories, for example, divide scenes into regions of common illumination (“frameworks”) but do not explicitly represent the illuminant within a shared field of illumination (Bressan, 2006; Gilchrist et al., 1999). Anchoring models apply normalization and scaling rules to recover surface albedo, which do not embody an explicit representation of the illuminant. Filter models also involve no representation of illuminants and surfaces; they simply generate a transformed brightness map of the retinal luminance distribution. One of the persisting problems in attempts to adjudicate theories of lightness (and brightness) is that it is often difficult to develop critical tests that will differentiate between different theories of the same phenomena. The checker–shadow illusion, for example, can be explained through concepts of frameworks and anchoring (Radonjic, Whyte, Faasse, & Gilchrist, 2006), linear filtering (Shapiro & Lu, 2011), or as a form of scission wherein the visual system discounts the different amounts of illumination. These models yield different expectations, if not explicit predictions, about whether similar effects should be obtained with chromatic versions of such displays. We consider how our results impact on these different models in turn.

First, consider what a scission model would (or at least could) predict. We have previously argued that the luminance or chromatic variations across contours and textures can induce a scission into multiple components (Anderson & Khang, 2010; Anderson & Winawer, 2005; Wollschläger & Anderson, 2009). If scission qua transparency is responsible for the induction observed in the checker–shadow illusion, then we would expect that chromatic and achromatic variants of this display would exhibit a similar pattern of induction. The checkerboard pattern provides a spatial regularity that could, at least in principle, provide evidence that the gradient was a secondary component of modulation and, hence, treated as a separate causal layer. Although this analysis could account for the achromatic stimuli, it cannot provide any insight into the failure to observe the same boost in induction in the chromatic versions of this display.

It is also unclear how (or whether) filtering models or local forms of contrast induction could explain the different pattern of results in the chromatic and achromatic variants of these displays. It is well known that comparable chromatic and achromatic induction effects can be observed in a variety of simple display types, including simultaneous contrast, Mach bands, and White’s effect. To the extent that such effects can be understood on the basis of low-level filtering processes, or contrast enhancement and filling-in, it would be reasonable to expect comparable effects in the checkerboard–gradient display if such factors played a significant role in
generating these displays. However, one persisting difficulty involved in making this comparison is the significant differences in the spatial and temporal sampling properties of chromatic and achromatic channels (Granger & Heurtley, 1973; van der Horst & Bouman, 1969). Chromatic sensitivity falls off more rapidly at high frequencies than luminance channels, whereas luminance channels are more bandpass (at least when measured in comparable ways). Although this makes it difficult to directly compare chromatic and achromatic displays, it seems unlikely that the differences in spatial sensitivity can account for the specific pattern of data we report herein, particularly if we consider both the tips–strips pattern and the checkerboard–gradient pattern. The check sizes for the checkerboard patterns and tips and strips pattern were essentially identical in our displays. The low-frequency component in the strips pattern was also the same as the gradient frequency in the checkerboard–gradient image. Nonetheless, the achromatic and chromatic tips and strips pattern exhibited a similar pattern of results, but the chromatic and achromatic checkerboard–gradient images did not. Moreover, the pattern of induction observed in the checkerboard-only pattern was in the direction of contrast, not assimilation, which would not be predicted from the lower cutoff frequency of chromatic channels. This component was also a larger percentage of the effect observed in the color displays than in the achromatic displays, so it would seem more likely for the chromatic checkerboard–gradient pattern to show an even larger boost in effect size than the achromatic pattern when the two components are combined. Overall, the data suggest that it is unlikely that the low-pass characteristic of chromatic channels can account for the difference we observed between the achromatic and chromatic checkerboard–gradient displays. In support of this view, we also informally compare the effects of viewing distance on the relative effect size of the different component patterns. We did not observe a discernable effect of viewing distance on the order of the effects for either the chromatic or achromatic patterns.

One of the assumptions that motivated our use of chromatic displays involves a presumed difference in how the visual system interprets local variations in chromaticity versus local variations in luminance. Chromatic boundaries can arise from changes in the spectral content of the illuminant, which occurs commonly in natural scenes containing shadows that receive their light from secondary reflections and/or diffuse illumination. Nonetheless, it is much more likely that strong variations in chromatic content arise from changes in the spectral reflectance of surfaces than from differences in the prevailing illumination. If the visual system is biased to interpret local chromatic changes as reflectance changes, then a display containing purely chromatic contrast should exhibit a bias to be perceived as changes in intrinsic surface color. This suggest that if the lightness difference observed in an achromatic display involves computations that attribute some of the luminance variation to differences in the illumination, it should be difficult to generate an equivalent chromatic variant of the display that exhibits the same pattern of induction. The general prediction is consistent with the pattern of data we obtained for the checkerboard–gradient pattern. This suggests that the induction observed in the checkerboard–gradient image may involve computations that are specific to estimates of illumination intensity rather than spectral content.

In conclusion, we have found that similar patterns of induction can be observed for chromatic and achromatic displays that evoke strong percepts of homogenous transparency, which suggests that scission may play a causal role in the induction observed in displays that induce percepts of transparency. In contradistinction, we observed different patterns of induction in displays containing chromatic and achromatic luminance gradients, which suggests that the large boost in induction observed in the achromatic checkerboard–gradient patterns is specific to luminance modulations. Any successful model of lightness or color must provide a cogent explanation for when such differences are obtained and when they are not.

Acknowledgments

This research was supported by grants from the Australian Research Council to B. Anderson.

Commercial relationships: none.
Corresponding author: Barton L. Anderson.
Email: barta@psych.usyd.edu.au.
Address: A19, Sydney, New South Wales 2006, Australia.

References


