Lightness, brightness, and anchoring

Barton L. Anderson
School of Psychology, University of Sydney,
Sydney, Australia

Michael Whitbread
School of Psychology, University of Sydney,
Sydney, Australia

Chamila de Silva
School of Psychology, University of New South Wales,
Sydney, Australia

The majority of work in lightness perception has evaluated the perception of lightness using flat, matte, two-dimensional surfaces. In such contexts, the amount of light reaching the eye contains a conflated mixture of the illuminant and surface lightness. A fundamental puzzle of lightness perception is understanding how it is possible to experience achromatic surfaces as specific achromatic shades in the face of this ambiguity. It has been argued that the perception of lightness in such contexts implies that the visual system imposes an "anchoring rule" whereby a specific relative luminance (the highest) serves as a fixed point in the mapping of image luminance onto the lightness scale ("white"). We conducted a series of experiments to explicitly test this assertion in contexts where this mapping seemed most unlikely—namely, low-contrast images viewed in dim illumination. Our results provide evidence that the computational ambiguity in mapping luminance onto lightness is reflected in perceptual experience. The perception of the highest luminance in a two-dimensional Mondrian display varied monotonically with its brightness, ranging from midgray to white. Similar scaling occurred for the lowest luminance and, by implication, all other luminance values. We conclude that the conflation between brightness and lightness in two-dimensional Mondrian displays is reflected in perception and find no support for the claim that any specific relative luminance value acts as a fixed anchor point in this mapping function.

Introduction

A basic problem in vision science involves understanding how the visual system recovers the reflectance properties of surfaces. One extensively studied aspect of surface reflectance is *lightness*, which is typically defined as the proportion of light a surface reflects diffusely. The reflectance properties of opaque surfaces can be characterized by their bidirectional reflectance distribution function (BRDF), which in general can be quite complex. However, models of surface reflectance typically decompose the BRDF into diffuse and specular components, corresponding to the lightness (or color) and gloss of a surface, respectively. The majority of lightness studies have focused on simple planar surface geometries with idealized Lambertian reflectance functions that scatter light uniformly in all directions (see, e.g., Annan & Gilchrist, 2004; Arend & Spehar, 1993a, 1993b; Bressan, 2006; Economou, Zdravkovic, & Gilchrist, 2007; Gilchrist et al., 1999; Gilchrist, 1977, 1979; Kingdom, 2011). One putative advantage of such stimuli is that they reduce the number of different sources of image structure that arise in natural scenes (e.g., shading, irradiance flow, specular reflections). However, this simplicity comes at an informational cost. The flat, matte surfaces that dominate work in lightness perception are arguably the most ambiguous lightness stimuli possible; diffuse surface reflectance is inextricably conflated with both the illuminant and the three-dimensional pose of a surface. Indeed, for a flat, matte, surface, any luminance can be generated by any possible reflectance by pairing it with an appropriate illuminant, and, moreover, any luminance can appear as any surface lightness when placed in an appropriate context.

Despite their inherent ambiguity, we typically experience surfaces as having a specific (or absolute) lightness. It is unclear how this percept is possible given the information available in the images. For surfaces with a common surface pose and illuminant, some information about relative lightness is theoretically available, but the same is not true for absolute lightness, particularly for the kind of images that have

Citation: Anderson, B. L., Whitbread, M., & de Silva, C. (2014). Lightness, brightness, and anchoring. *Journal of Vision, 14*(9):7, 1–13, http://www.journalofvision.org/content/14/9/7, doi:10.1167/14.9.7.

doi: 10.1167/14.9.7 Received March 27, 2014; published August 7, 2014 ISSN 1534-7362

dominated research in lightness. For surfaces with a common pose and illuminant, a fixed luminance ratio (say, 2:1) implies a fixed reflectance ratio of 2:1, which implies that luminance ratios could in principle provide sufficient information to specify relative reflectance. Note, however, that this ratio can be satisfied by an infinite number of different reflectance pairs (e.g., 20% and 40%, 8% and 4%, 15% and 7.5%, etc.), so some additional information—information that does not appear to be present in the images—is needed to generate a representation of absolute surface lightness. This computational ambiguity suggests that the perception of absolute lightness is in some sense imposed by the visual system, and has inspired attempts to discover how the visual system maps luminance onto perceived lightness in the face of this ambiguity.

Anchoring theory (Gilchrist et al., 1999) proposed a set of rules that were offered as a solution to the problem of mapping luminance onto perceived lightness. Anchoring theory decomposes the luminance–lightness mapping problem into two subproblems: a *scaling* problem and an *anchoring* problem. The scaling problem involves determining how differences (or ratios) of luminance are mapped onto lightness intervals. The anchoring problem specifies how these intervals are "affixed" to the lightness scale. The primary claim of anchoring theory is that a particular relative luminance value serves as a fixed reference point on the lightness scale (i.e., it acts as an "anchor"). Any number of anchoring points is theoretically possible. For example, the average luminance could be assigned a value of midgray, the highest luminance could be assigned a value of white, or the lowest luminance could be assigned a value of black. There has been no theoretical or physical rationale offered for selecting a particular luminance to treat as an anchor point. To date, this choice has been made solely on the basis of empirical data, where a variety of experiments have suggested that the visual system interprets the highest luminance as white (~90% reflectance).

Some of the most compelling evidence in support of anchoring theory arose from studying what were regarded as the "simplest" displays that support a percept of a surface. These stimuli ranged from a full-field stimulus divided by a simple edge (Li & Gilchrist, 1999) to two-dimensional Mondrian displays containing an array of different reflectances (Cataliotti & Gilchrist, 1995). The approach taken by Gilchrist and colleagues was to "…consider the rules of anchoring under minimum conditions for the perception of a surface and then to attempt to describe how the rules change as one moves systematically from simple images to complex images" (Gilchrist et al., 1999, p. 799). There are a number of issues that arise in attempting to understand the perception of lightness using such reduced stimuli. There is a host of information in natural scenes that could potentially be used to disentangle the contributions of the light source, surface reflectance, and surface pose that is missing in the simple images that have dominated research in lightness perception, including the stimuli that have shaped anchoring theory. For virtually all of the stimuli that shaped anchoring theory, the illuminant and reflectance of a surface are maximally conflated. Indeed, the distinction between lightness and brightness was arguably the most poorly defined in such stimuli since there is essentially no information available to disentangle the contribution of the illuminant from the reflectance of surfaces. Given the conflation between brightness and lightness in these displays, it would seem somewhat remarkable if this ambiguity were not also evident in perception.

In what follows, we present a series of experiments that were designed to experimentally assess whether the ambiguity between lightness and brightness is, in fact, also reflected in perception. To anticipate what follows, we report that the perception of lightness in two-dimensional Mondrian displays scales with both a surface's brightness and the contrast of the surrounds in which it is embedded, and find no evidence that the highest luminance (or any other) serves as a fixed anchor point in the luminance-to-lightness mapping function for low contrast and low luminance Mondrian displays.

Experiment 1

The purpose of Experiment 1 was to determine whether the highest luminance in a two-dimensional Mondrian appears white when viewed in a room where observers are immersed in the illumination field. Previous experiments that have provided support for anchoring have used displays that were either illuminated in a room by a hidden spotlight or inside boxes in a homogeneous and relatively bright illumination. In natural viewing contexts, observers typically have some awareness of the light field in the sense that they have expectations about the intensity and direction of light sources for objects inserted into a scene (Koenderink, Pont, van Doorn, Kappers, & Todd, 2007). In previous work on anchoring, this information was either completely missing or, in the case of hidden spotlights, explicitly violated. A second potential issue with previous work supporting anchoring is that the matching display used as the measurement device often depicted a sequence of calibrated papers placed on a white background. Essentially any theory of lightness or brightness would predict that the perceived lightness of the match targets would be shifted downwards when placed on a light surround, which would artificially

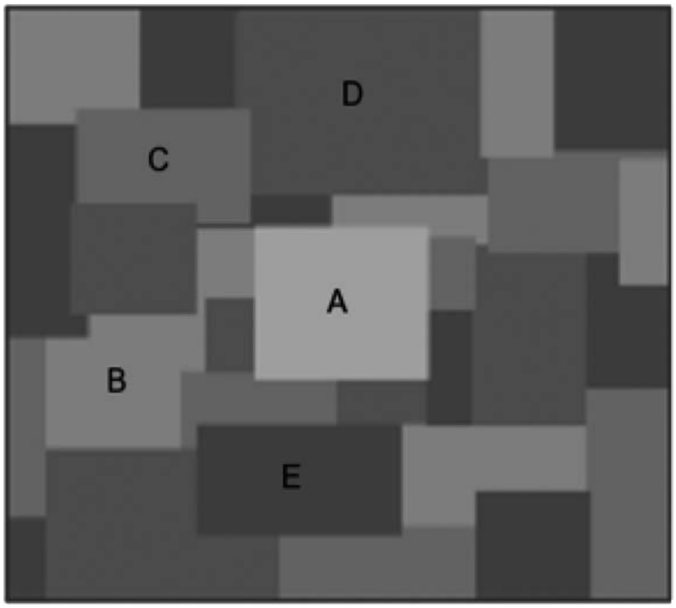

Figure 1. A schematic of the Mondrian displays used in Experiment 1. The central patch (marked A) was always the highest luminance and was bordered by each of the other four surface reflectances by equal amounts.

inflate observer's lightness matches (i.e., they would have to select a higher lightness to match the same shade of gray as something placed on a darker surround). The purpose of the first experiment was designed to determine (1) whether the highest luminance in a two-dimensional Mondrian would appear white when viewed in a black room illuminated by dimmable, overhead room lights; (2) whether there are any measurable effects of the background on which the matching targets are placed, using either a white or a random noise background; and (3) whether observers' lightness judgments are affected by the intensity of the illumination. We reasoned that it would be more likely to observe a failure of mapping the highest luminance to white for dark, low-contrast surfaces viewed in dim illumination.

Observers

Forty-eight members of the University of New South Wales (34 females and 14 males) served as observers. Thirty-five were first year students who received course credit for their participation, and the remaining were more advanced students or University employees who volunteered their time.

Stimuli

Five different Mondrians were constructed from calibrated Color-Aid (CA; Color-aid Corp, Hudson Falls, NY) papers that span a scale of 1 to 10 in steps of 0.5 (which are intended to be approximately equal perceptual intervals). Each Mondrian contained five consecutive gray scale values of CA papers, four of which served as the surround, with the lightest appearing only once in the center of the figure. The spatial configuration of the Mondrians was always the same, and is shown in Figure 1. Regions with the same shade in the figure represent regions in the Mondrians that were made from the same CA paper. The spatial arrangement of the patches was identical in each stimulus. The stimuli were designed such that the central target patch was bordered by each of the papers used in the background by an equal length.

The five stimuli were created using CA papers ranging between CA 1 (3.1% reflectance; black) and CA 5 (27.2% reflectance; middle gray, a Munsell value of ~5.75). The percentage reflectances were calculated by calibrating CA papers with the Munsell neutral (gray) scale. The experimenter and two additional judges compared the CA paper with Munsell paper to obtain the best possible approximation of the Munsell value and corresponding reflectance values.

Mondrian A (the lightest Mondrian) was constructed with CA papers CA 5 (27.2% reflectance), CA 4.5 (24.6%), CA 4 (22.1%), CA 3.5 (13.7%), and CA 3 (12%). Under the high illumination used in the experiment, these papers generated luminance values of 14, 12.4, 10.8, 7.5, and 6.4 cd/m^2, respectively. Mondrian B was constructed with CA papers CA 4.5, CA 4, CA 3.5, CA 3, and CA 2.5. The luminance of these papers under high illumination was 12.4, 10.8, 7.5, 6.4, and 5.0 cd/m^2, respectively. Mondrians C (CA 4, CA 3.5, CA 3, CA 2.5, and CA 2), D (CA 3.5, CA 3, CA 2.5, CA 2, and CA 1.5), and E (CA 3, CA 2.5, CA 2, CA 1.5, and CA 1) were all constructed in the same way (reflectances of remaining CA chips CA 2, CA 1.5, and CA 1 were 7.7%, 4.6%, and 3.1%, respectively). The range of reflectance ratios was 2.27:1, 2.37:1, 2.87:1, 2.98:1, and 3.87:1 for Mondrians A through E, respectively.

Each stimulus had the same design of 30 overlapping rectilinear shapes ranging in size from 7.62 × 7.62 cm to 15.24 × 12.7 cm. The center of each stimulus was a 10.16 × 10.16 cm square varying between CA 3 (12%) and CA 5 (27.2%) and was always the highest luminance in the stimulus. The top right corner was a 7.62 × 12.7 cm rectilinear shape and always had the lowest luminance in the stimulus, varying between CA 1 (3.1%) and CA 3 (12%).

Procedure

Observers were taken into a black room in which no visible surface was brighter than the brightest patch on

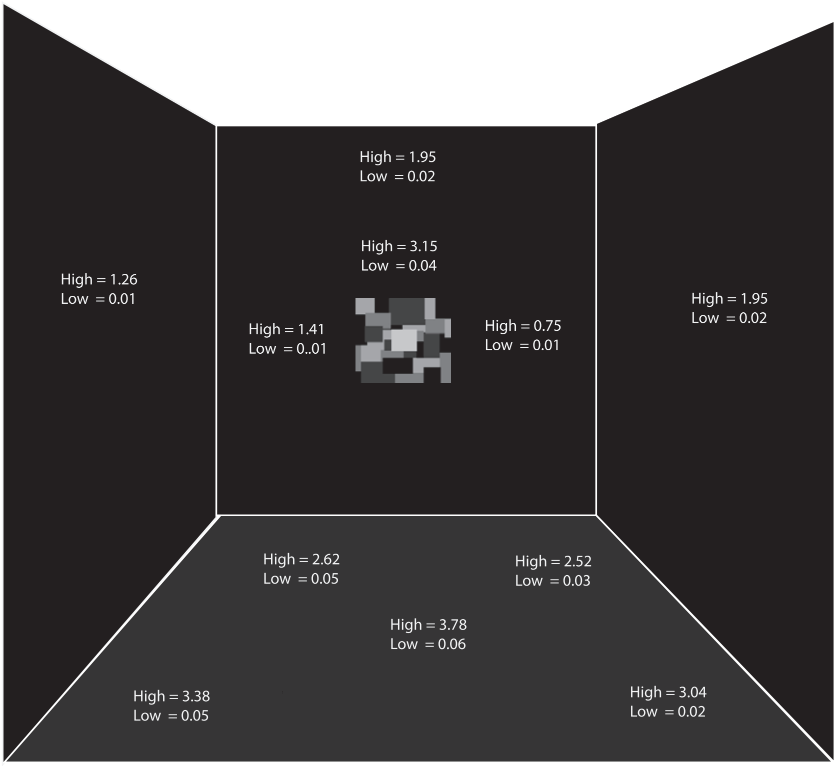

Figure 2. A diagram of the room used in Experiments 1 and 2. The two numbers indicate the luminance of the walls in the high- and low-illumination conditions used in the experiments.

the Mondrian. The room was illuminated by overhead adjustable florescent lights that were set to either their highest or lowest setting (corresponding to the high-illumination and low-illumination conditions) and were not within the field of view of observers. The amount of light reflected from the walls and floor was measured and is presented in Figure 2. The intensity of the light striking the target was approximately 51.75 cd/m^2 in the high-illumination condition and 0.754 cd/m^2 in the low-illumination condition. The highest luminance in Mondrians A through E was 14.1, 12.7, 11.43, 7.1, and 6.21 cd/m^2, respectively, in the high-illumination condition and 2.1, 1.9, 1.7, 1.0, and 0.9 cd/m^2, respectively, in the low-illumination condition.

Participants were led into the black room with their eyes closed and stood on a mark on the floor that was 1.5 m from the wall from which the Mondrians were viewed. The Mondrian subtended 15.2° and the central target patch subtended 3.87° of visual angle. Observers were required to keep their eyes closed while one of the Mondrians was placed on hooks on the wall, which were initially covered with a matte black foam board when observers entered the room. The experimenter then moved behind the observer, at which point observers were instructed to open their eyes. Observers viewed the Mondrian under either high illumination (lights on as bright as possible, ~51.75 cd/m^2) or low illumination (lights dimmed as far as possible, ~0.754 cd/m^2). The subjects then walked into an adjacent room illuminated by florescent lights (~64 cd/m^2) and selected one of the CA papers on one of the two matching backgrounds. Each target in Mondrians A through E was matched twice under both high and low illumination, where the repeat was rotated 180° relative to the initial viewing position. The order of the Mondrians and their orientation were both counterbalanced across observers. Matches were made using one of two CA scales: one where 19 CA papers (from CA 1 to CA 10 with increments of 0.5) were affixed to a white background or one where the same 19 papers were affixed to an articulated background (black and white random-dot surround). Each subject made 20 matches—two replications of all 10 conditions (five Mondrians and two illumination conditions).

Participants were randomly allocated to either the "white scale" or "articulated scale" condition. Each participant made 10 matches (two repetitions of each of the five Mondrians) in one illumination condition, then 10 matches (two repetitions of each of the five Mondrians) in the other illumination condition. The illumination condition used first was counterbalanced across participants. In the dim-illumination condition, participants were required to undertake 1.5 min of dark adaptation before making each match. During this time, participants were required to fixate the black walls during the adaptation period and then closed

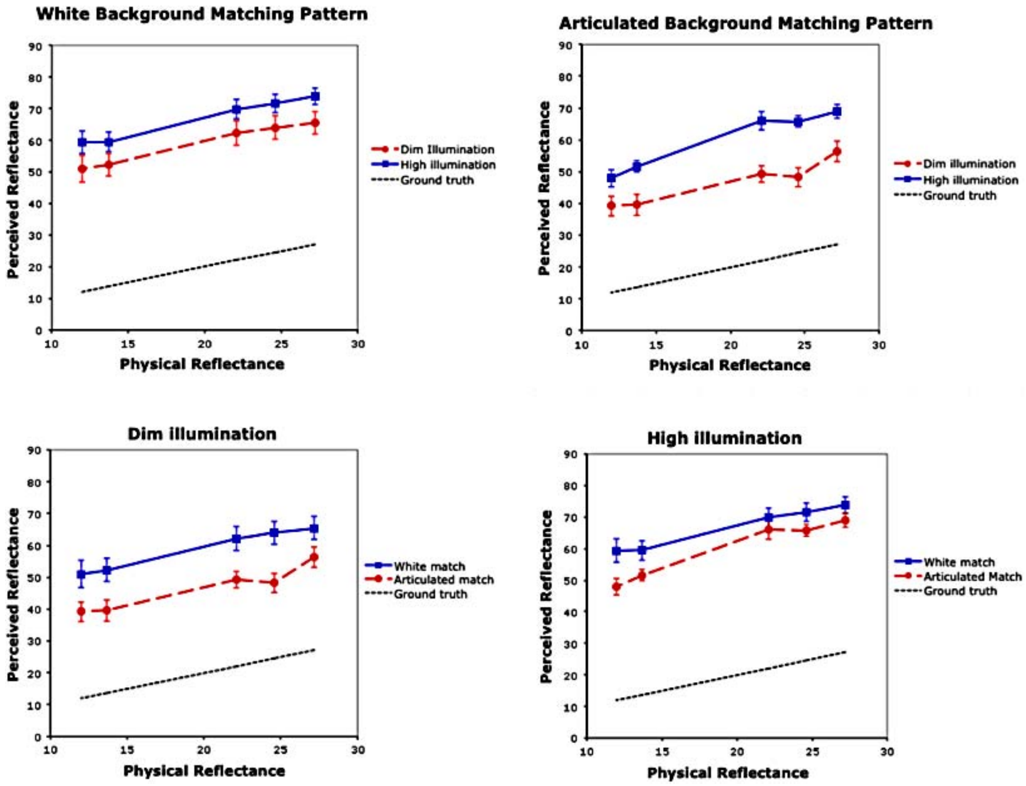

Figure 3. The results of Experiment 1. The top row presents the data for the two different matching backgrounds (white and articulated, respectively), whereas the bottom presents the data for the two different illumination conditions. Anchoring theory predicts that all points should have the same value (~90% reflectance).

their eyes as the experimenter placed one of the Mondrians on the wall. They then made their lightness judgments of the Mondrian.

Results

The results of Experiment 1 are presented in Figure 3. Each panel in the top row depicts the results for a fixed matching chart (white or articulated) for the two different illumination conditions (bright or dim), whereas each panel in the bottom row shows the same results for one illumination condition for the two matching charts. Each data point represents one of the five different Mondrians (E–A from left to right) viewed in one of the two illuminants and matched to one of the two matching charts. The veridical lightness of the targets is plotted as the dashed line. As is evident in all of the panels, the perceived reflectance of the highest luminance target increases as a function of the target's true reflectance value, none of which is perceived as white (~90% reflectance). Note that anchoring theory predicts that the highest luminance should be independent of the physical reflectance of the target patch, and all appear as the same reflectance (~90%). The perceived lightness of all of the targets is, however, much lighter than their true reflectance values (note the vertical offset from the dashed line). The perceived reflectance of the central targets is consistently higher in the high-illumination condition than in the low-illumination condition for both chart conditions ($p < 0.001$, binomial sign test[1]) and always matched to lighter matches for observers who used the CA chart on the white background than for those who used the CA chart on the articulated matching background ($p < 0.001$, binomial sign test).

The results of this experiment demonstrate that when Mondrians are viewed in an illumination field in which observers are immersed, they do not necessarily map the highest luminance onto "white" on the lightness scale. Indeed, there is no evidence that there is any fixed anchor point in these data; the perceived lightness of the highest luminance scaled with its true reflectance (and hence brightness) and varied as a function of the illumination level (which also increased brightness). Note that the different Mondrians all spanned a different range of reflectances, so the effect of increasing the surface reflectances within the Mondrians also increased the contrast of the Mondrian relative to the wall. The effect of illumination level, however, cannot be similarly understood, as the change in illumination affected both the Mondrian and the walls on which it was placed. Finally, we found that the background of the matching chart had a substantial effect on observers' lightness matches. Matches to the chart on the white background were consistently higher than the same matches made with the chart with the articulated surround, which suggests that previous work using matching charts with white backgrounds may have overestimated the perceived lightness experienced by observers.

Experiment 2: Computer Mondrians

The results of the previous experiment revealed that the perception of lightness of the most reflective surface (highest luminance) varied as a function of its true reflectance, the illumination level (brightness), and the type of background used for the match pattern. The purpose of Experiment 2 was to investigate the effects of illumination level more thoroughly in a more controlled experimental setting. In this study, we generated computer variants of the Mondrians used in Experiment 1 and varied the illumination over a much broader range of values than those used in Experiment 1. We confirmed in a control experiment that the same pattern of results obtained in Experiment 1 was observed with a computer-simulated version of the same experiment, which allowed us to explore the effects of simulated illumination in a more controlled setting. The goal was to assess how perceived lightness of the highest luminance varied as a function of the Mondrian contrast and simulated illumination level (brightness) for a broad range of changes in simulated illumination.

Participants

Thirty-seven first-year psychology students from the University of New South Wales participated in Experiment 2 in exchange for course credit.

Stimuli

Two new Mondrians (A and B) were constructed. They had the same geometric pattern as those used in Experiments 1 and 2 but had different simulated reflectance ranges. Mondrian A had a simulated reflectance range of 4:1 and Mondrian B had a simulated reflectance range of 1.7:1. They were constructed so that the highest simulated reflectance was the same in the two Mondrians (6%), and the other patches were selected such that they were approximately equally spaced in reflectance; the lowest luminance generated a ratio of either 4:1 (A) or 1.7:1 (B). The reflectances for A were (6, 4, 3, 2, 1.5) and for B were (6, 5.23, 4.66, 4, 3.5).

Thirteen different simulated illumination levels were created for each of the Mondrians (1, 8, 15, 90, 180, 270, 380, 480, 600, 700, 800, 930, and 1030 cd/m^2), which were selected to span the range of luminances available to the monitor when multiplied by the chosen reflectances. The reflectances where chosen by calibrating a monitor (Lacie Electron Blue IV, LaCie, Tigard, OR) and finding five values that generated luminance ratios of either 4:1 or 1.7:1 and that had the same highest simulated reflectance (6%). These reflectances were then multiplied by 12 different illumination factors to generate luminances (in cd/m^2) with the range of the monitor. The 12 new Mondrians so derived were constructed by finding lookup table values that produce luminances (in cd/m^2) that matched the luminance derived from multiplying the simulated reflectance by the simulated illuminant. Twenty-four Mondrians were created in this manner [2 reflectance ranges (Mondrians A and B) × 12 simulated illuminants], which resulted in 26 stimuli in the experiment [the 24 derived and the two original (illumination = 1) Mondrians]. The brightness of the central (target) square is the same in both Mondrians A and B for any given simulated illuminant.

Procedure

The procedure for Experiment 2 is similar to that for Experiment 1. The experiment was conducted in the room depicted in Figure 2, and both the room and monitor were illuminated using the low-illumination condition used in Experiment 1. The presentation order of the 26 Mondrians was randomized, and each subject matched the central (target) square in each of the 26 Mondrians twice (52 trials) using only the articulated scale from Experiment 1.

Participants were asked to walk into the room displayed in Figure 2 and sat in a chair 1.5 m from the computer screen. The stimuli subtended an angle of 15.2° (width) by 11.2° (height). They were instructed to close their eyes for approximately 10 s, during which a new display was generated by the experimenter. Observers were allowed to look at each Mondrian as long as necessary to gain a clear impression of the lightness of the central patch (the highest luminance). They then walked out of this room into an adjacent room, where they selected the match from the same articulated background display used in Experiment 1. The order of the displays was randomized within an experimental session and across observers.

Results

The perceived lightness of the highest luminance in the Mondrian is plotted in Figure 4. The data reveal a

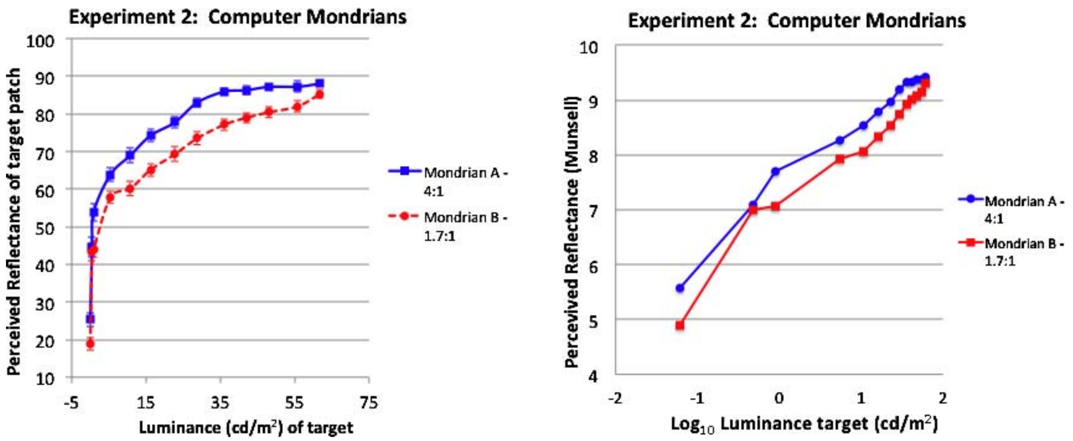

Figure 4. The results of Experiment 2. The left graph depicts the results in terms of perceived reflectance as a function of luminance, whereas the right graph presents the data in terms of perceived Munsell values as a function of log luminance.

strong dependence of perceived lightness on the simulated illumination, or brightness, of the Mondrian. This dependence is characterized by compressive nonlinear growth in perceived lightness as a function of luminance, which is reasonably captured by a logarithmic function (r^2 values of 0.99 and 0.97 for Mondrians A and B, respectively). The data also reveal a consistent and significant effect of the contrast of the Mondrian: The perceived lightness of the target in the higher contrast (4:1) Mondrian is consistently lighter than the same luminance embedded in the lower contrast (1.7:1) Mondrian ($p < 0.0001$, binomial sign test). The most significant violations of the highest luminance rule occurred for Mondrian B for the lowest simulated illuminant, which appeared to have a reflectance of 18.8% (a Munsell value of 4.89); the same luminance in Mondrian A appeared to have a reflectance of 25.35% (a Munsell value of 5.57; Munsell values derived from Newhall, Nickerson, & Judd, 1943). The data do not provide support for the claim that the highest luminance serves as a fixed anchor point on the lightness scale, which predicts that perceived lightness should be independent of the simulated illumination level or luminance. Consistent with Experiment 1, these results demonstrate that the perceived lightness of the highest luminance increases as a monotonic function of its brightness: The highest luminance appeared midgray for the lowest luminance values and increased in perceived lightness over the entire range of simulated illuminants (and luminance values) tested.

Experiment 3

The results of Experiments 1 and 2 reveal that the perceived lightness of the highest luminance in a two-dimensional Mondrian display varies as a function of its brightness and the contrast of the Mondrian in which it is embedded. To this point, we have measured the effects of luminance and contrast for only the highest luminance in the Mondrian, so it is unclear how or whether the perceived lightness of the other surfaces varies as a function of simulated illumination (brightness). In Experiment 4, we created two new Mondrians (C and D) with simulated reflectance ratios of 1.3:1 and 25:1, respectively, and had observers match both the lightest and darkest patches in the Mondrian. These two contrast ranges were chosen to evaluate the generality of our effects for a different range of luminance ratios and to assess whether lightness scaling is affected by either brightness or contrast.

Participants

Twenty-six first-year psychology students from the University of Sydney participated in exchange for course credit.

Stimuli

The stimuli were created in the same manner as in Experiment 2. The stimuli were designed such that the highest luminance would be the same for the two Mondrians (C and D) but satisfied two different ranges of reflectance ratios (~1.3:1 or ~25:1). The simulated reflectances of Mondrians C and D were (0.06, 0.0573, 0.0537, 0.05, 0.0467) and (0.06, 0.0326, 0.017, 0.0067, 0.0024), respectively. It was not possible to produce versions of Mondrians C and D in the darkest two illumination conditions (illuminant = 1 and 8) because of a lack of sufficient luminance range in the monitor. Hence, the highest luminance was 0.9 cd/m^2 in the lowest simulated illumination condition

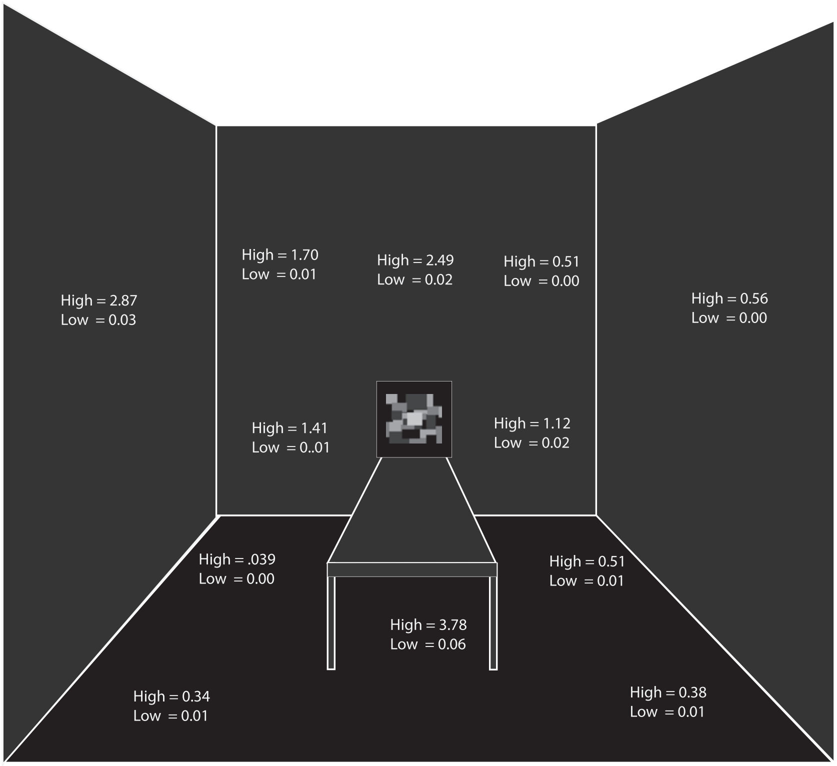

Figure 5. The experimental room where Experiments 3 and 4 were performed, with the corresponding luminances for the high- and low-illumination conditions of different surfaces in the room.

that would satisfy the 25:1 ratio, unlike the 0.6 cd/m^2 value in the simulated illuminant of 1 in Experiment 3. By design, the 11 simulated illuminants of Mondrians C and D have the same central test patch luminance as the brightest 11 versions of Mondrians A and B, but the darkest test patch in Mondrian C was always brighter by a factor of ~19:1 to the darkest patch in Mondrian D.

Procedure

The procedure in Experiment 3 is essentially identical to that in Experiment 2, with two main differences. Experiment 3 was conducted in a lab at the University of Sydney rather than the University of New South Wales. The rooms were very similar in design and size. Both rooms were painted matte black, and the carpet in the room at the University of Sydney was constructed out of black velvet. A luminance map of the room at the University of Sydney under the experimental illumination conditions is presented in Figure 5. For this experiment, the stimuli were viewed in the dim-illumination condition to ensure that the highest luminance in the Mondrian display was always the highest luminance in the scene.

The procedure, monitor, and viewing conditions were identical to those in Experiment 2, with one significant change. Observers were required to make two matches: one to the highest luminance (central patch) and one to the lowest luminance (upper right corner). Observers first viewed the stimulus of the central patch (highest luminance) until they felt they had a clear impression of its lightness, and then walked into the main lab room where they viewed the same CA matching chart used in Experiment 2. They then returned to the dimmed room, returned to their seat, closed their eyes for 10 s, and viewed the same Mondrian and were asked to establish an impression of the lightness of the darkest patch (of the upper right corner). Subjects then again left the room to perform a lightness match. Each subject made 88 matches [2 Mondrians (C and D) × 11 simulated illumination conditions × 2 repetitions × 2 targets (darkest and lightest)], which were all randomized across trials and observers.

Results

The results of Experiment 3 are presented in Figure 6. As with the results of Experiment 2, the perceived lightness of the highest luminance exhibited a compressive nonlinear growth over the range of simulated illuminants tested, which were well fit with logarithmic functions (r^2 of 0.99 and 0.97 for Mondrians C and D,

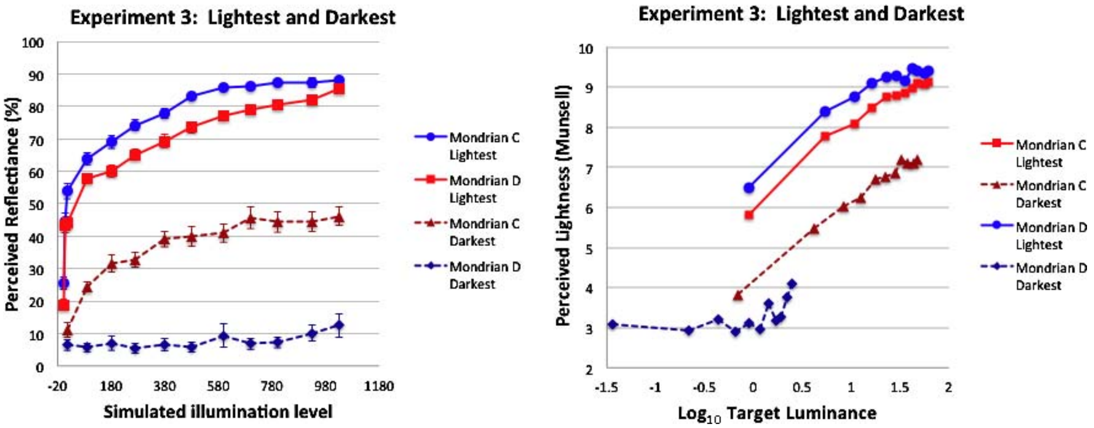

Figure 6. The results of Experiment 3. The left plot presents perceived reflectance as a function of simulated illumination level, whereas the right plot presents perceived lightness in Munsell values as a function of the log target luminance. Note that the highest luminances in the two Mondrians were identical but the lowest luminances were not.

respectively). We also observed a statistically reliable difference in the perceived lightness of the target for the two Mondrian contrasts (binomial sign test, $p < 0.0004$). As with Experiment 2, the lightness of the highest luminance was always lighter in the higher contrast Mondrian than in the lower contrast Mondrian, although the effect is relatively modest in magnitude (~0.7 Munsell steps over the range of luminances tested). The darkest square in Mondrian C also exhibits the same compressive nonlinear increase as the lightest square, which parallels the increase in the lightness observed with the lightest square. This is most readily appreciated when the perceived lightness is plotted as Munsell values against log luminance of the test patches: The two highest luminances, and the lowest luminance in the low-contrast Mondrian, appear to increase in lightness at the same rate. In contrast, the darkest square in the higher contrast (25:1) Mondrian is largely unaffected by the change in simulated illuminant and appears close to black for all simulated illuminants tested.

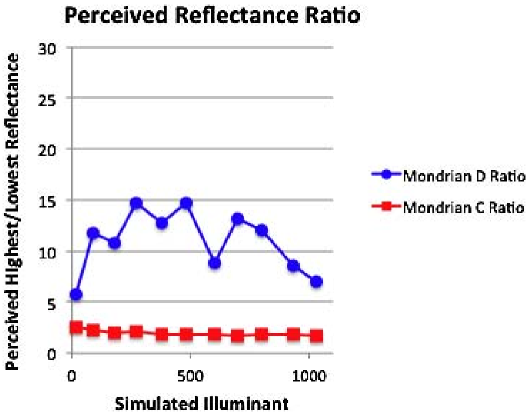

Figure 7. The perceived reflectance ratios of the Mondrians in Experiment 3. The range perceived in C was modestly higher than its true ratio, whereas Mondrian D's ratio of 25:1 was highly compressed.

Figure 7 plots the scaling data of the two Mondrians as a ratio of the highest luminance to the lowest luminance. Mondrian C had a 1.3:1 reflectance range, which was modestly expanded and perceived as spanning a range from 2.6:1 to 1.8:1, decreasing monotonically as a function of simulated illuminant (or luminance). In contrast, Mondrian D had a 25:1 reflectance range, which was significantly compressed, being perceived as having as little as a ~6:1 range for the dimmest luminance, with a maximum range of ~15:1 for intermediate luminance values. Thus, although theoretically there was sufficient information to recover the relative reflectance values of the lightness in both Mondrians, this information is apparently not used in the manner suggested by Wallach's ratio principle (see also Arend & Spehar, 1993a, 1993b). The simulated illuminants used in this experiment were explicitly designed to preserve luminance ratios, yet the perceived range of reflectances varied significantly (particularly for the Mondrian with the 25:1 reflectance range).

Experiment 4: Temporal anchoring control

Our final experiment sought to determine whether our results could have arisen, at least in part, from some form of temporal anchoring. This experiment was designed to check whether the failure to perceive the highest luminance as white was a consequence of being exposed to higher luminances immediately prior to the Mondrians observers judged, such as occurred when viewing the matching charts in the other room. Some previous work has provided some evidence for such

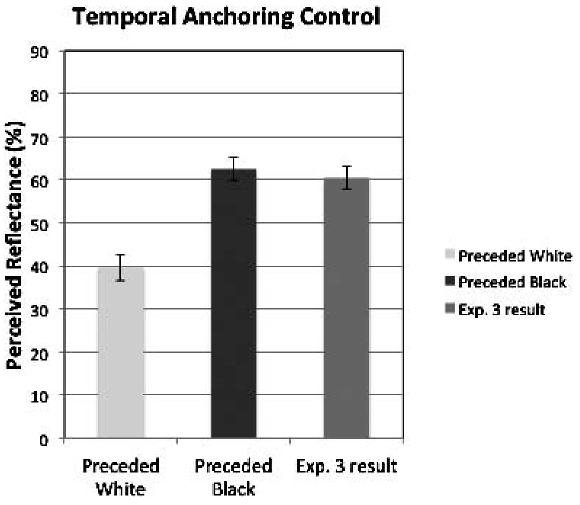

Figure 8. The results of the temporal anchoring control experiment, together with the same condition of Experiment 3. There was no evidence for temporal anchoring affecting lightness matches in the conditions reported in Experiments 1, 2, and 3; such effects were observed only when immediately preceded by a white patch in the experimental monitor.

temporal effects. Cataliotti and Bonato (2003) found that that such temporal effects disappeared by 32 s, whereas Annan and Gilchrist (2004) observed temporal effects that lasted up to 2 min. We performed a number of informal tests to evaluate whether such effects could account for our failure to observe the highest luminance as white by dark adapting for protracted periods of time in our black room before viewing the display. Although we observed no substantial effects informally, we designed a control experiment to assess whether such temporal effects could account for the kind of scaling we observed in our experiments by having observers judge one of our target patches after 2 min of fixating a completely dark CRT screen or while performing a distraction task that involved exposure to white.

Observers

Twenty-six observers from Experiment 3 participated in Experiment 4 at the conclusion of the 88 matches they made in Experiment 3. Thirteen observers each were assigned to one of two groups described below.

Stimuli and procedure

Experiment 4 took place in the same room as Experiment 3. The observers were split into two groups, which we refer to as "preceded black" and "preceded white." For the "preceded black" condition, observers viewed a dark computer screen depicting a homogenous black field for 120 s. This was the same screen used to display stimuli in Experiments 2 and 3. After a 120-s interval, Mondrian C (1.3:1 simulated reflectance range) from Experiment 3 was presented at a simulated illuminant of 180, and observers were asked to match the central (brightest) square of the Mondrian (which had a luminance of 10.8 cd/m^2). Once the stimulus was displayed on the computer screen, observers left the room and again matched the lightness of the central square using the articulated CA scale used in the previous experiments.

In the "preceded white" condition, observers performed a lightness-matching task based on a simultaneous contrast stimulus. During these 120 s, subjects were required to adjust the luminance of a square patch by means of a computer mouse to match a series of different targets on different colored backgrounds. The background of the display contained a random-dot pattern that contained a mixture of the highest luminance the monitor could output (61.8 cd/m^2) and the lowest ($\sim$0 cd/m^2), and in some of the trials both the target and the backgrounds of the simultaneous contrast display also had the maximal luminance (experienced as white). At the conclusion of these 120 s, the Mondrian stimulus was displayed on the screen, and observers left the room and matched the lightness of the central square using the articulated CA scale. There was no gap or break between the presentation of the distraction task and the presentation of the stimuli. The stimulus was presented immediately at the end of 120 s of distraction task.

Results

The results of Experiment 4, together with the match made to the same target in Experiment 3, are presented in Figure 8. The match to the target made in the condition preceded by a dark screen was statistically indistinguishable from the matches made to the same target in Experiment 3 ($t = 0.77$, $p > 0.05$). There was, however, a substantial effect ($t = 8.06$, $p < 0.001$) of immediately preceding the Mondrian display with a distraction task in which the highest luminance available on the monitor was presented (which appeared white). In this case, the perceived reflectance of the target is reduced by $\sim$24%, which is a difference of $\sim$1.75 Munsell steps). Thus, although we can observe effects of temporal anchoring when a Mondrian is immediately preceded by stimuli that contain higher luminances that appear white, we do not observe any such effects for the conditions in which we conducted Experiments 2 and 3.

General discussion

The primary purpose of the experiments presented herein was to determine whether the conflation of lightness and the illuminant in two-dimensional Mondrian displays is reflected in perception. More specifically, we assessed whether the highest luminance is mapped onto a fixed surface reflectance ("white") independently of target reflectance, illumination level (real or simulated), or luminance. Our data reveal that brightness is perceptually conflated with lightness over the range of luminance values we tested. It is known that the "highest luminance rule" of anchoring theory fails in contexts where observers experience a luminous light source. However, the data presented here reveal a much more general failing of the concept of anchoring. We find no evidence to support the claim that the highest luminance is treated as a fixed anchor point in the luminance to lightness mapping, but rather, the perceived lightness of the highest luminance scales with both its brightness (simulated illuminant) and its contrast. We are forced to conclude that the highest luminance does not play the role of a lightness anchor. Rather, it is best understood as an approximate mapping of the highest luminance onto the lightness scale that applies to only a limited range of luminance values, contrasts, and geometric contexts.

Although we did not find evidence in support of anchoring, we did observe that the reflectance of the highest luminance was consistently overestimated in Experiment 1, where the ground truth of reflectance had some physical meaning. This overestimation may, in fact, be a consequence of the flat, matte surfaces that have dominated studies on lightness perception. Most natural surfaces contain three-dimensional shape information on a variety of scales: a global three-dimensional shape, medium-scale surface relief (mesostructure), and the microstructure responsible for the scattering of light. Although the visibility of the global shape and the invisibility of the microstructure responsible for the scattering of light are essentially independent from a surface's reflectance, the same would not be true for many forms of mesostructure. High-reflectance surfaces will increase the local luminance and diminish the intensity of luminance gradients arising from local surface relief by the scattered light of interreflections, both of which would reduce the contrast generated by the mesostructure. The overestimation of surface reflectance of flat, matte surfaces may arise, at least in part, because the visual system treats the *absence* of visible mesostructure as a cue that the reflectance of a surface is relatively high. Some data in support of this conjecture were reported for surfaces that contained clearly visible mesostructure. Sharan, Li, Motoyoshi, Nishida, and Adelson (2008) investigated the perception of lightness of a broad range of surfaces that contained visible mesostructure and varied in albedo. They took calibrated photographs of these surfaces and then equated their mean luminance. Their data were more consistent with a "gray world" hypothesis than with anchoring the highest luminance to white: Light surfaces were judged to be physically darker than they were, and dark surfaces appeared lighter, with regression slopes that passed through midgray.

The need for a lightness anchor has been justified by the empirical observation that we experience surfaces as possessing a specific, absolute reflectance (albedo) despite the fact that the images contain information that can only, at best, specify relative surface lightness. This has led to a theoretical assertion that the perception of absolute reflectance requires some means by which at least one *particular* relative luminance is anchored to an absolute lightness scale (Gilchrist et al., 1999). Despite the central role this assertion has played in anchoring theory, it is not logically entailed by the experience of (absolute) reflectance. The experience of absolute lightness implies only that the visual system has some means of mapping luminance (or some other measure of image structure) onto an absolute lightness scale; it does not imply that any particular relative luminance serves as a fixed anchor point. Indeed, the results of the experiments presented herein provide strong evidence against the concept of a fixed anchor point. If there is any validity to the concept of an anchor in shaping the data presented herein, it is more appropriately construed as a floating anchor that shifts as the intensity of the image luminance (and contrast) varies.

The focus on the anchoring problem was also shaped by the belief that relative lightness values could, at least in principle, be recovered from the images. This idea was formalized by Wallach (1948), who argued that there is a one-to-one mapping between relative luminance and relative lightness (e.g., a 2:1 ratio in luminance would be mapped onto a 2:1 ratio in perceived lightness). For surfaces embedded in a common illuminant, this rule makes physical sense since variations in the intensity of the illuminant will leave the ratios of luminances in the images unchanged and hence provide veridical information about relative surface reflectance that is invariant to changes in illumination. Empirically, however, this simple ratio principle can fail in a variety of ways (see, e.g., Arend & Spehar, 1993a, 1993b). Two of the simplest forms of failure occur when the perceived range of reflectances exceeds the physical range of reflectances (gamut expansion) or when the perceived range of reflectances is less than the physical range (gamut compression). Both forms of rescaling have been observed in the literature, and both are evident in the results of Experiment 3 described herein. Particularly dramatic failures of lightness scaling have been reported in experiments with displays that span a much broader range of luminance ratios than can be theoretically obtained by reflectances in a common illuminant. Radonjic, Allred, Gilchrist, and Brainard (2011) reported that a stimulus ratio of 5905:1 could be mapped onto an extended lightness ratio of 100:1 and concluded that such results ruled out theories that predict perceived lightness from luminance ratios or Weber contrast. Indeed, these data cast significant doubt on the view that the visual system has any understanding of the range of reflectance values that populate natural environments.

The studies presented herein reveal that the visual system conflates lightness with brightness in simple, two-dimensional displays of the kind that have dominated a great deal of research into lightness perception. It is often claimed that lightness constancy is quite good in natural scenes (see, e.g., Kingdom, 2011), but often little data are offered in support of this view. Robilotto and Zaidi (2004) attempted to assess this claim by having observers perform lightness dissimilarity comparisons of natural surfaces in different illuminants. Two pairs of achromatic objects (cups covered with folded and crinkled paper) were presented in two different illuminants and immersed in a highly articulated scene containing a full range of surface reflectances. One of the cups had a different reflectance than the other three. Observers were required to find the cup that differed in reflectance across the change in illuminants. Observers were generally quite poor at this task and appeared to rely on brightness dissimilarity rather than explicit estimates of surface lightness. Thus, lightness constancy can be quite poor, even in scenes that contain a broad range of reflectances, three-dimensional shape cues, and interreflections.

In conclusion, the data presented here provide strong evidence against the view that the visual system treats a single relative luminance as an invariant anchor point in the mapping from image luminance to lightness. The ambiguity between brightness and lightness in two-dimensional Mondrian displays is also evident in perception, and our experiment of lightness varies as both a function of luminance and contrast, at least for the range of luminance values and contrasts explored herein.

Keywords: lightness, surface perception, anchoring, color constancy

Acknowledgments

The research reported herein was supported by grants from the Australian Research Council to B. Anderson.

Commercial relationships: none.

Corresponding author: Barton Anderson.
Email: barton.anderson@sydney.edu.au.
Address: University of Sydney, Sydney, Australia.

Footnote

[1] All of these effects reported as significant with nonparametric tests were also significant with parametric tests (ANOVAs); the nonparametric tests are reported for simplicity.

References

Annan, V., Jr., & Gilchrist, A. (2004). Lightness depends on immediately prior experience. *Perception and Psychophysics, 66*(6), 943–952.

Arend, L. E., & Spehar, B. (1993a). Lightness, brightness, and brightness contrast: 1. Illuminance variation. *Perception and Psychophysics, 54*(4), 446–456.

Arend, L. E., & Spehar, B. (1993b). Lightness, brightness, and brightness contrast: 2. Reflectance variation. *Perception and Psychophysics, 54*(4), 457–468.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review, 113*(3), 526–553, doi:10.1037/0033-295x.113.3.526.

Cataliotti, J., & Bonato, F. (2003). Spatial and temporal lightness anchoring. *Visual Cognition, 10*(5), 621–635.

Cataliotti, J., & Gilchrist, A. (1995). Local and global processes in surface lightness perception. *Perception and Psychophysics, 57*(2), 125–135.

Economou, E., Zdravkovic, S., & Gilchrist, A. (2007). Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *Journal of Vision, 7*(12):2, 1–15, http://www.journalofvision.org/content/7/12/2, doi:10.1167/7.12.2. [PubMed] [Article]

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X. J., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*(4), 795–834.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science, 195*(4274), 185–187.

Gilchrist, A. L. (1979). The perception of surface blacks and whites. *Scientific American, 240*(3), 112–122, 124.

Kingdom, F. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research, 51*(7), 652–673, doi:10.1016/j.visres.2010.09.012.

Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M., & Todd, J. T. (2007). The visual light field. *Perception*, *36*(11), 1595–1610.

Li, X., & Gilchrist, A. L. (1999). Relative area and relative luminance combine to anchor surface lightness values. *Perception and Psychophysics, 61*(5), 771–785.

Newhall, S. M., Nickerson, D., & Judd, D. B. (1943). Final report of the O.S.A. subcommittee on the spacing of the Munsell colors. *Journal of the Optical Society of America*, *33*(7), 385–418. doi:10.1364/JOSA.33.000385.

Radonjic, A., Allred, S. R., Gilchrist, A. L., & Brainard, D. H. (2011). The dynamic range of human lightness perception. *Current Biology, 21*(22), 1931–1936, doi:10.1016/j.cub.2011.10.013.

Robilotto, R., & Zaidi, Q. (2004). Limits of lightness identification for real objects under natural viewing conditions. *Journal of Vision, 4*(9):9, 779–797, http://www.journalofvision.org/content/4/9/9, doi:10.1167/4.9.9. [PubMed] [Article]

Sharan, L., Li, Y. Z., Motoyoshi, I., Nishida, S., & Adelson, E. H. (2008). Image statistics for surface reflectance perception. *Journal of the Optical Society of America A: Optics Image Science and Vision, 25*(4), 846–865.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology, 38*(3), 310–324.