Local and global processes in surface lightness perception

JOSEPH CATALIOTTI
Upsala College, East Orange, New Jersey

and

ALAN GILCHRIST
Rutgers University, Newark, New Jersey

Various demonstrations show that a target of constant luminance can be made to appear darker in perceived lightness merely by introducing an adjacent region of higher luminance. This has often been interpreted as a manifestation of contrast effects produced by lateral inhibition, a relatively local process. An alternative interpretation holds that the highest luminance in such a display serves as an anchor that defines the white level. This interpretation is global in the sense that the anchor need not be located near any particular target in order to serve as its standard. Edge integration processes have been postulated that would enable such remote comparisons, but there is controversy about the strength of these processes. We report a series of experiments in which local and global processes were assessed. Specifically, we tested whether the introduction of a higher luminance has a greater darkening effect on an adjacent target than on a remote target. We found no difference, suggesting that the darkening effect is a matter of anchoring, not contrast, and that edge integration processes required by anchoring are relatively strong.

One of the most remarkable demonstrations in the field of black and white perception is the Gelb effect, described by Koffka (1935, p. 245) as follows:

> In a dark room a perfectly homogeneous black disk is rotated; this disk, and nothing else, is strongly illuminated by a projection lantern. Under these conditions the disk looks white and the room black. Then the experimenter holds a small piece of white paper close to and in front of the rotating disk so that it falls within the cone of light. At the same moment the disk alters its appearance, and looks black.

Wallach (1948), inspired by Gelb's work, demonstrated that a homogeneous disk of light that appears white or luminous can be made to darken, perceptually, to any shade of gray without changing its intensity (called luminance) of an annulus of light that surrounds it. In both of these demonstrations, a target surface darkens perceptually when the luminance of an adjacent region is increased. Why does this happen? There are two answers in the literature on lightness perception. In short, one answer involves contrast and the other involves anchoring.

1. The contrast answer: According to some models of lightness perception (Cornsweet, 1970; Heinemann, 1955; Jameson & Hurvich, 1964; Leibowitz, Mote, &

Thurlow, 1953), the crucial mechanism in lightness perception is lateral inhibition among neural units. By this account, the disk in the demonstrations of Gelb or Wallach appears to darken because, although the intensity of light traveling from the disk to the corresponding retinal receptors does not change, there is increased inhibition acting on those receptors from adjacent or nearby receptors stimulated by light from the adjacent part of the visual field.

2. The anchoring answer: According to a different approach, the highest luminance in a scene tends to appear white and serves as a standard by which the lightness value of other surfaces is judged. When the highest luminance increases, the standard is raised and other surfaces appear darker, not because they have changed, but because they are now darker with respect to the highest luminance. Something close to this account can be found in Koffka (1935), Wallach (1948), Land and McCann (1971), and Horn (1986).

The anchoring concept is so fundamentally important to the perception of surface lightness that it is difficult to understand why it has received so little explicit attention. No model of lightness perception can predict a specific value of perceived lightness for any surface without an anchoring rule. For example, the ratio principle (Wallach, 1948), despite its importance, cannot by itself yield specific lightness values. Recall that in his classic experiments, Wallach (1948) demonstrated only that targets having equal luminance ratios appear equal in lightness. Nothing was said about the specific lightness values. Relative luminance values per se can only produce

relative lightness values. Specific or absolute lightness values require an anchoring rule in addition to the ratio principle.

Consider a concrete example: a disk/annulus pattern in which the luminance of the annulus is five times greater than that of the disk. Wallach, though he gave it very little attention, assumed an anchoring rule whereby the highest luminance appears white. This in turn implies that the disk will appear middle gray. The annulus appears white because it is the highest luminance, and the disk appears middle gray because the reflectance of middle gray is five times less than the reflectance of white, just as the luminance of the disk is five times less than the luminance of the annulus. But other anchoring rules are equally compatible with the ratio principle. The rule could be that the lowest luminance appears black. In that case, the disk would appear black (because it has the lowest luminance), not middle gray, and the annulus would then appear middle gray. Another possible rule, one actually given by Helson (1964), is that the average luminance appears middle gray. This rule would predict that the annulus would appear light gray and the disk would appear dark gray. Yet other rules are also possible (Gilchrist & Bonato, in press). The highest luminance rule is the one that has been most often invoked (Horn, 1986; Koffka, 1935; Land & McCann, 1971; Wallach, 1976), especially for disk/annulus and Mondrian displays, and indeed empirical results (Bruno, 1992; Gilchrist & Jacobsen, 1983) suggest that it is generally valid for such displays. Yet this rule cannot apply in all cases, or else the perception of luminous regions would be impossible (Bonato & Gilchrist, in press).

Note that the highest luminance rule implies a visual capacity not usually associated with a contrast mechanism—the capacity to compare the luminance of regions that are quite remote from each other in the visual field. As Land and McCann (1971) have pointed out, the visual system does not have the ability to insert a standard luminance next to any target surface in a display. But they and others (Arend, 1973; Arend, Beuhler, & Lockhead, 1971; Gilchrist, Delman, & Jacobsen, 1983; Whittle & Challands, 1969) have postulated a process that would achieve the same thing: edge integration. According to Land and McCann, the visual system integrates chains of ratios, and each ratio is picked up by a very closely spaced pair of photoreceptors.

Reid and Shapley (1988; see also Shapley & Reid, 1985; Shapley, 1985) have correctly pointed out that Land and McCann's edge integration mechanism, as far as it has been described, fails to account for local contrast effects. For example, a gray surface that happens to be completely surrounded by a white region in a Mondrian will appear darker than another surface of equal gray that is surrounded by black. This problem led Reid and Shapley to propose a weaker model of edge integration, according to which brightness (which is equivalent to lightness for a set of adjacent coplanar surfaces without an obvious illumination boundary) is mainly dependent on the local contrast at the border of a target surface. The effect of local contrast is only partially antagonized by edge integration (which they call *assimilation*), and the effect of edge integration is held to weaken with retinal distance. Reid and Shapley claim that the strength of edge integration is about 60% of the strength of local contrast, even at short range (5′ of visual angle), and that it weakens to a relative strength of 34% at 43′ of visual angle. These measurements were obtained by using a display in which one disk/annulus pattern was placed on a dark surround and an identical disk/annulus pattern was placed on an adjacent lighter surround. The lighter and darker surrounds were separated by a luminance ramp in one instance and a luminance step in another. In essence, Reid and Shapley measured the proportion of the simultaneous contrast effect of the light and dark surrounds on the two annuli that was transferred to the two disks. They found that, on average, the disks appeared to differ in lightness only about 50% as much as the annuli appeared to differ (from each other) in lightness. They reasoned that if the lightness of the disks was determined solely by local contrast, the disks would appear equal because they have equal local contrast. On the other hand, if edge integration could fully antagonize the effects of local contrast, the disks would appear to differ just as much as the annuli, because the effects that the light and dark backgrounds exert on their respective annuli would be fully passed on (through edge integration) to the disks.

Reid and Shapley's (1988) critique of Land and McCann is analogous to the issue of whether the darkening of the Gelb disk should be attributed to processes of contrast or anchoring. The commonality concerns the difficult question of local versus global processes in lightness perception. In Reid and Shapley's model, as in contrast models based on lateral inhibition, the major emphasis is placed on local interactions. Conversely, models that feature anchoring by the highest luminance imply that interactions between remote regions of the visual field can be just as effective, or at least nearly as effective, in setting lightness values as local comparisons.

We have conducted a series of experiments that address this controversy. Our tests were built on the classic Gelb (1929) effect under conditions to which the highest luminance rule would apply. Using black paper as the target surface, we tested whether it was darkened more when the highest luminance was adjacent to it than when other surfaces were spatially interpolated between the two. This allows a test of the strong versus the weak versions of edge integration and a test of the anchoring interpretation of the Gelb effect versus the contrast interpretation. If target lightness is the same in both cases—that is, if the visual system can compare luminance values as readily when the two surfaces are separated by other surfaces as when the two are adjacent—it supports the anchoring interpretation and a strong form of edge integration. But if the target appears significantly darker when it is adjacent to the highest luminance than when it is remote, then this supports a strong emphasis on the role of local contrast.

In Experiment 1, rather than using the simple traditional method involving two surfaces, one black and one white, we successively built up an adjacent row of five squares. Beginning with a single black square, we then added an adjacent dark gray square. We next added a middle gray square adjacent to the dark gray one. In the same manner we added a light gray square, and finally a white square. Each square, as it is added, becomes the new highest luminance and should darken the previous squares, including the original black square. The question is whether the darkening power of these squares on the black square weakens, or is just as great, as their displacement from the black square increases.

In Experiment 2 we introduced a between-subjects design to control for experience effects, and we introduced configurations that broke the linkage between number of targets in the display and the distance between target and standard. In Experiment 3, we varied the retinal distance between the target and the highest luminance merely by changing the viewing distance of the observer. Finally, in Experiment 4 we compared adjacent versus remote darkening in a chamber we called a "Mondrian world" in order to eliminate from the visual field the outer framework that consisted of the dimly illuminated laboratory.

EXPERIMENT 1

Method

Laboratory arrangements. Between one and five test squares were presented under Gelb illumination conditions—that is, in the beam of a projected cone of bright illumination within an otherwise dimly illuminated laboratory. The test squares were 10-cm (visual angle of 1.5º) squares of achromatic Color-Aid paper suspended in midair and supported by a set of five aluminum rods (80 cm long, 1.27 cm thick), each of which was hidden by the test square it supported. Test square reflectance values were approximately 3%, 12%, 30%, 60%, and 90%. The arrangements are shown in Figure 1. The special illumination for the test squares was provided by a Kodak Ektagraphic B-2 projector located 2.75 m diagonally in front of and to the right of the squares, and at the same height as the squares. The luminance of the black square in the projector beam was 1.4 ftL. The projector beam was cropped, by means of a metal slide containing a rectangular aperture, so that it was just large enough to illuminate the entire row of five squares. The projector was hidden from the observer by a black cloth suspended next to it. The part of the projector beam that overshot the test squares and illuminated a section of the back wall of the laboratory was hidden from the observer by a baffle. The observer sat in a chair at a viewing distance of 3.7 m, directly in front of the test squares.

General room illumination came from a 500-W incandescent flood lamp housed in a special chamber located near the ceiling, behind the observer. A 16 × 18 cm square opening in the front of the chamber housing was covered with a translucent plastic panel for diffusion. Horizontal strips of heat-resistant Roscolux colored transparent material were attached to this panel in order to equate the color of this light source to that of the projector beam. The intensity of illumination was controlled by sliding a square opaque panel across the opening in the front of the chamber. The general room illumination was adjusted so that a black test square in the projector beam appeared completely identical to a true white test square located just outside the projector beam. The walls and contents of the laboratory were easily visible, though dimly illuminated.

Observers indicated the perceived lightness of these squares by selecting a match from a separately illuminated Munsell chart of 16 samples mounted on a rectangular piece of white cardboard. The chart was illuminated by a 15-W fluorescent tube mounted 10 cm above the chart so that the luminance of the whitest white (Munsell 9.5) was 160 ftL. The chart was located immediately in front of and below the observer.

Design. Five groups of 10 observers served. The first group saw only the black square, the second saw the black square plus the dark gray square, and so on, until the fifth group saw the entire row of five squares.
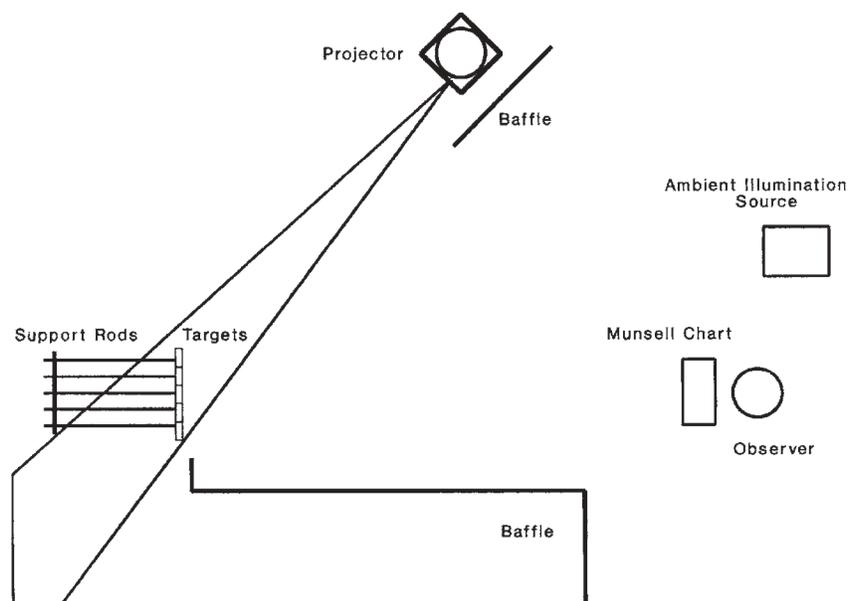

Figure 1. Plan view of laboratory arrangements.

Procedure. After being seated, an observer was shown the black target and was asked to give a verbal description of the test square and then make a match using the chart. Then the observer was asked to look away while the second, dark gray square was added. Both squares were then matched and the process continued until all five squares were exposed. No restrictions were made on head or eye movements during the matching. Following the Munsell matches on each trial, half of the observers (5 out of each group of 10) were asked to judge whether or not each target surface appeared luminous. This measure was taken because of a concern that our simple Munsell chart might involve a ceiling effect.

Observers. All the observers volunteered to participate—some to fill a course requirement—and all were naive with respect to the purpose of the experiment.

Results and Discussion

The Munsell matches are shown graphically in Figure 2A; for comparison, the theoretically expected results are shown in Figure 2B. The mean values of perceived lightness are plotted as a function of the luminance ratio between each target surface and the highest luminance in the display. It is obvious that the obtained results can be seen as a vertically compressed version of the expected results. There are two implications in this pattern; the darkening effect of adding higher luminance values to the display (1) does not weaken with distance, and (2) is slightly less than half of what would be expected. We will discuss these in turn.

1. The strong version of edge integration, by itself, does not require any particular degree of darkening, only that the darkening for a remote anchor be approximately as great as that for an adjacent anchor. Support for a local contrast model and/or for the weak version of edge integration would show up in Figure 2A as a positively accelerating negative slope. The straight lines that predominate in the data indicate that a given increase in the ratio between the highest luminance and the target luminance causes a fixed amount of darkening of the perceived lightness of lower luminances, regardless of the distance between the target and the highest luminance. This is true for each of the targets.

2. The limited amount of darkening, illustrated in Figure 2A by the vertical compression, was a surprise to us. On the face of it, this suggests a serious failure of the ratio principle, according to which the range of perceived lightness values (more concretely, the luminance range of the Munsell matches) is at least roughly equal to the range of luminance values in the stimulus. We conducted a second experiment to rule out two possible explanations of this limited darkening. The first potential explanation involves the within-subjects design of Experiment 1. When observers see the same surfaces in the same environment on multiple trials, there is a tendency (we call it a *hysteresis effect*) to continue to see the surfaces as being the same lightness value as they appeared in the initial trial, even though the experimental manipulation should produce a different appearance (and would for an inexperienced observer). In many of Gilchrist's experiments (Gilchrist, 1980; Gilchrist et al., 1983; Gilchrist & Jacobsen, 1984), for example, a between-subjects design was used because it was found that the

effects are cut roughly in half if the same observers serve in both conditions of the experiment. Indeed, Newson (1958) reported just such a result working with the Gelb effect: "If the surface was first seen as white, the introduction of an inducing surface did not usually cause subjects to see the surface as such a dark grey as when it was first seen as grey" (p. 88). To eliminate this possible factor, a between-subjects design was used in Experiment 2.

The second potential explanation involves the possibility that the darkening was somehow reduced by the number of targets used, or by the relatively small luminance steps between them. Because the Gelb effect is often conducted by placing a single white surface next to the black target, we added this simple condition to Experiment 2.

Finally, in a more direct test of the effect of an adjacent versus a remote anchor, we added a condition in
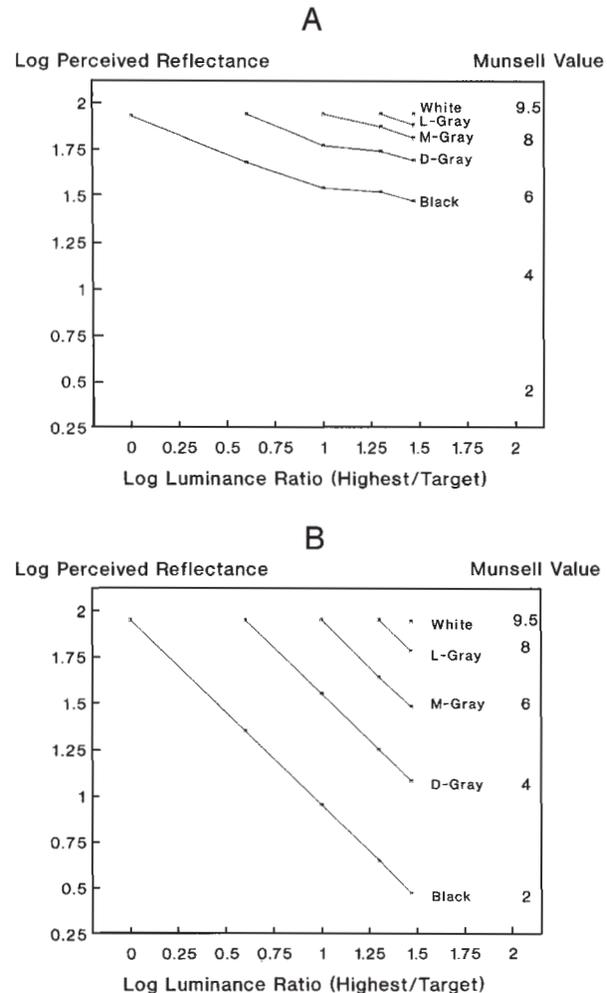

Figure 2. Perceived lightness as a function of luminance ratio between target and the highest luminance in the configuration. Each column of data points represents one configuration; each line connects data points for a single target (labeled by physical gray shade; L, light; M, middle; D, dark). (A) Obtained results; (B) expected results.

which the three middle-luminance targets were shifted one position to the right so that the white target could be inserted in the position adjacent to the black target. This allowed a comparison between the effect of a remote and an adjacent anchor without any change in the number of squares.

The luminosity judgments will be discussed together with the results of Experiment 2.

EXPERIMENT 2

Method

Seven different displays, shown in Figure 3, were each viewed by a separate group of 10 observers. Each of the first five groups viewed one of the five displays used in Experiment l. Another group was shown the black square with an adjacent white square and no other targets, and one group was shown the entire row of five squares with the white square adjacent to the black. In all other respects, the method was the same as that of Experiment 1.

Results and Discussion

Figure 4 shows the results obtained between subjects for the same conditions that were tested within subjects in Experiment 1. Apart from some additional noise (to be expected with the between-subjects method), these results essentially duplicate those obtained in Experi-

ment 1, showing that the lack of darkening cannot be attributed to a hysteresis effect. These results again show no distance effect. The lack of a distance effect is further underscored by the fact that with an adjacent white square, the black target appeared as a Munsell 6.0, exactly the same as the Munsell 6.0 with a remote white square. The distance function of a contrast model would clearly predict a difference between these two conditions, as would the distance function in the weak version of edge integration.

In the two-square condition, representing a closer replication of the Gelb effect, the darkening of the black square was not greater than that in the five-squares condition; it was actually less. The white square was matched to a Munsell 9.4 and the black square was matched to a Munsell 7.2, a value that is just barely significantly lighter [$t(18) = -2.11$, $p < .05$] than the 6.0 obtained in the five-squares condition. This indicates that the darkening failure cannot be attributed to the number of squares, nor to the size of the luminance steps.

The failure of the black square to appear darker than light gray is especially surprising. The Gelb effect is one of the most discussed effects in the field of lightness perception. Countless descriptions of the effect, beginning with early authorities such as Koffka (1935) and Katz
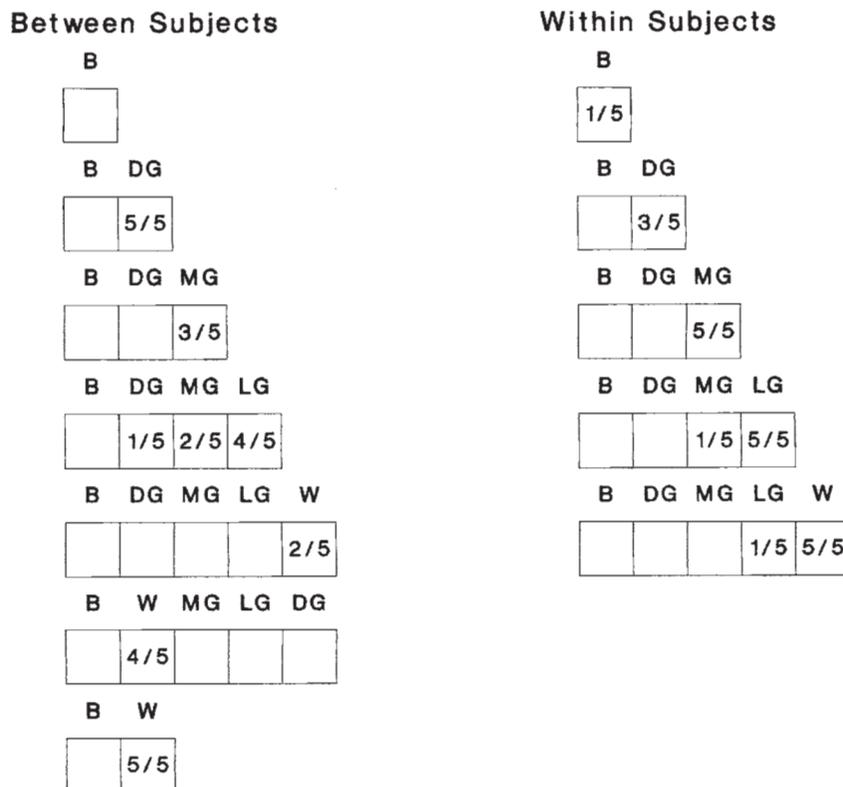
Luminosity Reports

Figure 3. Configurations used in Experiment 2, showing proportion of observers that reported each target as luminous. (B, black; DG, dark gray; MG, middle gray; LG, light gray; W, white.)
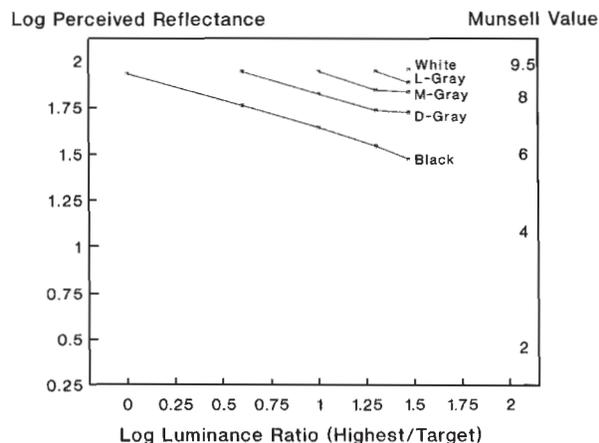
Figure 4. Results obtained between subjects in Experiment 2 for the same conditions that were tested within subjects in Experiment 1. (L, light; M, middle; D, dark.)

(1935), and including such prominent figures as Woodworth (1938), Wallach (1963), Evans (1948), and Helson (1964), have asserted that the black paper appears black when a white paper is placed adjacent to it. In three additional replications of the basic Gelb effect, as parts of additional experiments not reported in this paper, our observers matched the black paper (with an adjacent white paper) to Munsell values of 6.3, 7.3, and 7.6. Upon checking the literature, we discovered that this result is not peculiar to our laboratory. Despite the many informal descriptions of the Gelb effect, there have been only a few studies in which lightness values have actually been measured. Furthermore, those few studies were interested primarily in relative lightness values, not absolute lightness values, and none gave results in Munsell values. When we converted those various measures into Munsell values, we found data similar to our own results. Stewart (1959), in the most systematic study of the Gelb effect, varied the size of the white paper and its location (from the center of the Gelb disk to its periphery). Depending on conditions, he obtained lightnesses for the Gelb disk that ranged from 6.9 to 8.5 in Munsell values. Newson (1958) obtained values of 5.5, 4.9, and 5.3. Gogel and Mershon (1969) obtained a value of about 4.75. And McCann and Savoy (1991) obtained a lightness value of about 4.0, only by making the adjacent surface 300 times as intense as the black surface as well as much larger. Yet none of these investigators noted the inconsistency between their data and the common wisdom that the Gelb paper appears black under these conditions. Ironically, even Stewart repeated Gelb's and Koffka's reports that the disk turns completely black, but he failed to point out that his own results came nowhere near black.

At any rate, the failure of an adjacent white surface to darken the black Gelb paper to a perceived value below light gray has been a remarkably well-kept secret. The result is that the range of perceived grays in the target squares is compressed relative to the range of their phys-

ical grays. For example, in Experiment 1, a physical array of squares spanning the entire reflectance range from white to black (log luminance range, 1.48) was perceived with lightness values spanning less than a third of that range (log luminance range, 0.4). Although this issue is not central to the topic of this paper, and although we expect to report additional experiments that will address the causes of the compression, perhaps a brief discussion of this issue would be helpful here.

We know that the compression in the lightness range of the target squares is caused by what we call the outer framework: the surrounding, dimly illuminated environment. In Experiment 4, we showed that when the outer framework is eliminated, the range of lightness values in the inner framework is no longer compressed. One question is whether the apparent compression reflects a ceiling effect caused by the fact that our Munsell matching chart did not allow matches brighter than white—that is, self-luminous. Figure 3 shows that some of the targets were called luminous. But after extensive analysis we have rejected this hypothesis. The pattern of data obtained in Experiments l and 2 appears consistent with a vertical compression, not with a ceiling effect. A ceiling effect would produce a pattern of data in which the curve for each target would have two segments: a horizontal segment at the beginning representing the imposition of the ceiling, and a segment of slope -1 for the part of the curve below the ceiling. But a look at Figure 2A shows no such pattern. The data points fall on straight lines, indicating that all of the targets were affected equally, not simply those that fall above the ceiling.

This in turn raises the question of whether the compression represents a failure of the visual system to encode luminance ratios accurately at input, or whether the compression is the result of computations such as compromise or weighting. We believe the answer is the latter and will present evidence for this in a subsequent report.

EXPERIMENT 3

The weakening of the effectiveness of edge integration with distance hypothesized by Reid and Shapley (1988) appears to refer to retinal distance (visual angle). The spatial function generally attributed to lateral inhibitory interactions clearly refers to retinal distance. A very simple way to change the retinal distance in a display is to change the distance of the observer from the display. In Experiment 3 we presented the five squares at two viewing distances: at the same 3.7-m distance used in the previous experiments, and at only one third this distance. The near viewing position increased by a factor of three the retinal dimensions in all the displays. According to either a lateral inhibition model or Reid and Shapley's model, the darkening effect on the black square of a nonadjacent white square should be significantly reduced. According to the strong version of edge integration, the results should be similar to those obtained in the previous experiments.
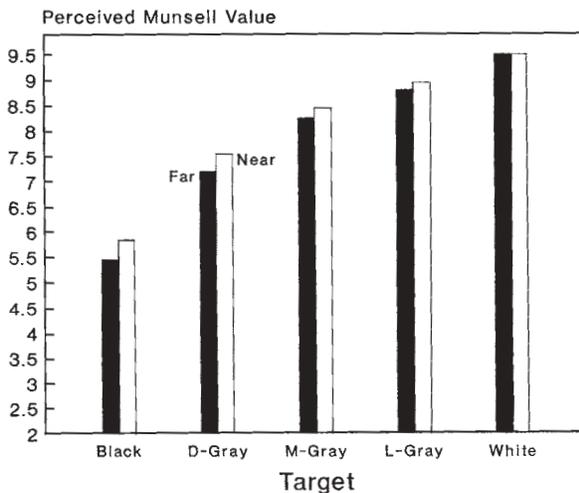
Perceived Munsell Value

Figure 5. Comparison of perceived lightness for each target in the far and near conditions.

Method

Ten observers viewed the five squares at the far distance, as before, and 10 viewed the five squares from a distance of 1.2 m. Otherwise the method was consistent with that of the previous experiments.

Results and Discussion

The results for this experiment are shown in Figure 5. None of the comparisons for corresponding targets yielded a significant difference. Although none of these differences were significant, they have a consistent direction. This probably indicates that there is a tiny effect in the direction that Reid and Shapley describe, but one that falls far short of their claims.

EXPERIMENT 4

In the experiments reported so far, we tested edge integration under conditions of multiple illumination frameworks—specifically, a region of special high illumination surrounded by an environment of low illumination. As we discovered, the surrounding field produces a perceived lightness range in the inner framework that is substantially compressed relative to its luminance range. This reduced ratio effect did not prevent us from comparing the addition of a higher luminance in an adjacent position with that of one in a remote position, but it did reduce by one half the effect of adding a higher luminance to the group, which was our vehicle for measuring the strength of edge integration.

In a final test of local versus global, we sought to eliminate the surrounding framework entirely, and thus its compressing effect on lightness values in the inner framework. In essence, we achieved this by proliferating the target squares until they completely crowded the surrounding framework out of the visual field. We created what might be called a "Mondrian world," a trapezoid-shaped room that completely filled the observer's visual field, covered entirely by a Mondrian pattern. We first tested lightness perception in such a room containing a limited range of dark grays. We expected that under these conditions the highest luminance, which in this case was a very dark gray, would be perceived as white. We then introduced a single, true white surface into the room and again tested perceived lightness. We were interested in the degree of darkening, produced by the addition of the white surface, for surfaces that were either adjacent to the white or remote from it. According to the weak version of edge integration, or according to a lateral inhibition account, the darkening effect should be strongest for adjacent surfaces, dropping in strength for surfaces farther and farther from the white. According to the strong version of edge integration, distance from the new higher luminance should matter very little.

Method

A chamber was constructed from foam board in the shape of a truncated pyramid resting sideways on one long edge of the base, as is shown in Figure 6. The large base served as the open side of the chamber, into which the observer's head was inserted. The entire inside of the chamber was covered with a collage of randomly placed rectilinear pieces of achromatic Color-Aid paper with reflectance values of 3%, 6.5%, and 12%, producing a total luminance range of only 4:1 within the chamber. The chamber was 76 cm deep and 100 cm wide at its base. Illumination was provided by two 200-W clear incandescent bulbs positioned on the left and right sides of the open side. The specific slant of the walls and the position of the light bulbs were chosen in order to prevent any stepwise difference in incident illumination at the corners. Voltage to the light bulbs was regulated so that a surface of 3% reflectance directly ahead of the observer had a luminance of 1.4 ftL, equal to that used in the previous experiments. A chinrest was mounted 15 cm into the chamber so that the observer's head was centered within the chamber.

Procedure. The observers were trained to memorize the Munsell scale using a set of 16 flashcards, each containing a 2.5-cm square of Munsell paper. Training continued until the observers were able to recognize the entire set to within half a Munsell unit.

After training, each observer was led into the completely darkened laboratory containing the Mondrian chamber. The observer was guided to the chinrest, which was covered with a luminous material. Once positioned, the observer was asked to keep hands to the sides and the head positioned forward. The illumination was then turned on and the observer was told the following: "Directly in front of you there is a square wall containing 26 overlapping surfaces. Your task will be to judge the color of each of the surfaces based on the scale you previously learned." Time was then allowed for questions and clarification. Particular targets were specified by use of shape and position descriptions—for example, "the C-shaped surface in the upper left-hand corner." Surfaces were judged in random order until all 26 target surfaces in the frontal plane were judged.

Design. Ten observers made judgments in the chamber, as described above ("before" condition), and a second group of 10 observers made judgments after a white target surface (90% reflectance), slightly larger in size than the average of the others, had been placed near the center of the frontal surface ("after" condition).

Results and Discussion

In the before condition, the surface with the highest luminance was perceived as a Munsell 8.9, even though it was physically a dark gray (Munsell 4.0), as would be
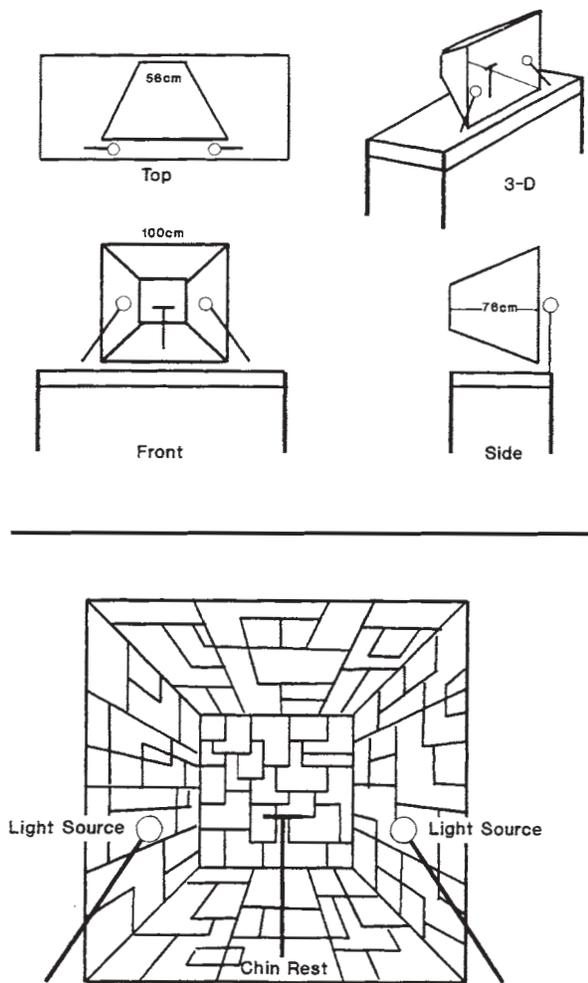
Figure 6. Apparatus used for the Mondrian world experiment.

age luminance rule, no white should be perceived under these conditions, and the highest luminance should be displaced away from white as far as the lowest luminance is displaced away from black. But although the perceived value of the highest luminance was only 0.6 Munsell steps away from an ideal white, the perceived value of the lowest luminance (Munsell 4.0) was two full Munsell steps away from black. These results would appear to favor the highest luminance rule, although this is somewhat obscured by the range expansion we found. We believe that a further test conducted with an even more restricted reflectance range would favor the highest luminance rule even more clearly.

Gilchrist and Jacobsen (1983) inadvertently conducted a test of this kind that was not confounded by a range expansion. They covered a two-dimensional Mondrian with a veiling luminance and presented it within a dark environment. The physical Mondrian contained the entire range of grays from black to white, but because the luminance of the veil was equal to that of a white in the Mondrian, the luminance range of the resulting display was reduced to only about 2:1. Observers (who did not perceive the veil itself) perceived the highest luminance as a Munsell 9.4 and the lowest luminance as a Munsell 6.6, just a bit lighter than middle gray. Those results favor the highest luminance rule over the average luminance rule quite decisively.

The mean lightness match for the Munsell 2.0 targets was 4.0 in the before condition and 3.7 in the after condition, a significant difference ($p = .016$). The mean lightness match for the Munsell 3.0 targets was 6.6 in the before condition and 6.3 in the after condition, a significant difference ($p = .015$). The mean lightness match for the Munsell 4.0 targets was 8.9 in the before condition and 7.3 in the after condition, a significant difference ($p < .0001$). For each of these three levels of reflectance, the adjacent surfaces did not darken significantly more than the remote surfaces. Adjacent surfaces were defined as surfaces that shared a border with the white square; remote surfaces were defined as surfaces that did not share a border with the white.

GENERAL DISCUSSION

We found support for the strong version of edge integration in four different kinds of tests.

1. When we successively introduced surfaces of higher luminance into the bright beam of light illuminating a square of black paper, the degree of perceptual darkening of the black paper depended only on the luminance relationship between the black paper and the highest luminance in the group, regardless of the distance between the black paper and the highest luminance.

2. When we rearranged the five squares so that the highest luminance was adjacent to the black square, instead of at the opposite end of the row, it made no difference.

3. The perceived lightness of the black square (and of the others as well) was the same whether the observer

expected by the highest luminance rule. What was not expected, however, was a substantial failure of the ratio principle, but this time we obtained an expansion, as opposed to the compression obtained in Experiments 1 and 2. The 6.5:1 range of perceived lightness values obtained in the before condition (log luminance range, 0.81) was much larger than the 4:1 range of physical luminance values (log luminance range, 0.60).

It might be noted incidentally that this condition by itself provides a test between the highest luminance rule and another popular rule, the average-luminance-as-middle-gray rule. The latter rule, closely related to what is sometimes called *the gray world hypothesis*, is attributed to Helson's adaptation-level theory. According to Helson (1964), any surface with a luminance equal to a weighted average of all the luminances in the visual field appears middle gray. The lightness values of other luminance levels are determined relative to this average. The restricted reflectance range of our before condition offers a clear test between the average luminance rule and the highest luminance rule. According to the aver-

viewed the display from a distance of 3.7 m or 1.2 m, even though the distance in visual angle between the black square and the highest luminance was three times as great when viewed from the near distance.

4. When a white surface was introduced into a Mondrian world populated only by surfaces ranging from black to dark gray, it lowered the perceived lightness value of all of the surfaces, regardless of whether they were adjacent to, or remote from, the white paper.

These results provide substantial support for the strong version of edge integration. They are also consistent with the high degree of constancy of surface lightness that characterizes our everyday experience. If the weak version were correct, surface grays would be seen correctly only if a white surface were adjacent or very close by in the retinal image.

Despite our failure, in all these tests, to find significant differences where Reid and Shapley would predict them, we did rather consistently find small, nonsignificant effects in the direction they would expect. So it would probably not be correct to conclude that retinal proximity has no effect whatsoever. But we suggest that Reid and Shapley's (1988) claim that edge integration antagonizes only about half of the effects of local contrast does not apply in the general case; it is probably applicable only to the rather unique conditions under which they tested it.

One of the features of Reid and Shapley's (1988) test is that the maximum change of lightness that can be produced is limited to the size of the simultaneous contrast effect. That effect is less than one Munsell step (Gilchrist, 1988), even for the standard pattern using black and white backgrounds that differ by at least seven Munsell steps. But it could be only a fraction of a Munsell step, given that Reid and Shapley's backgrounds differed by the equivalent of only one Munsell step. Such a limited change in lightness can produce a large ratio between the strength of edge integration and the strength of local contrast, even when the difference between the two strengths is quite small. Gilchrist (1988, Figure 4D) tested a configuration almost identical to that used by Reid and Shapley, except that he used black and white backgrounds. His data showed that the lightness of the two centers differed by more than that of the two surrounds, just the opposite of what Reid and Shapley report. But neither outcome should be treated as terribly reliable because when the maximum effect is so sharply limited to begin with, chance factors in the experiment can exert a disproportionate effect on the data.

McCann and Savoy (1991) attribute Reid and Shapley's (1988) results to a different feature of their conditions: restricted visual angle. McCann and Savoy measured the perceived lightness of a 2.5º test square at various degrees of separation from a much brighter region that took several shapes. They found a dramatic change, apparently due to the effect of scattered light from the brighter region, in target lightness at separations of between 0º and 1.25º. Beyond 1.25º, as the target escapes the stray light, further separation between the

target and the brighter region has only a negligible effect. McCann and Savoy point to the fact that Reid and Shapley measured integration across only 0.7º (well within the 1.25º region) as the key to understanding their results.

When the strength of edge integration is tested at greater separations and under conditions in which lightness can vary across the entire black/white range, integration is found to be quite robust.

For example, the well-known simultaneous contrast effect cited by Reid and Shapley (1988) as evidence of the failure of edge integration nevertheless illustrates its strength. The display consists of a gray target on a white background next to an equal gray target on a black background. To determine whether the appearance of this display supports the strong or weak version of edge integration, it is necessary to make a quantitative analysis. Specifically, one must take the obtained lightness difference between the two targets and see how this difference compares both with the difference that would be expected if edge integration had zero strength and with the difference that would be expected if edge integration had a strength of 100%. We can easily determine the difference that would be expected if the strength of edge integration were 100%; the targets would not appear to differ at all. And we can determine the obtained difference between the targets by consulting published reports of empirical results. Such reports (Berman & Leibowitz, 1965; Burgh & Grindley, 1962; Gilchrist, 1988; McCann, 1987; Mikesell & Bentley, 1930) indicate that in the simultaneous contrast display, the two targets appear to differ in lightness by between 0.5 and 1 Munsell steps. But it is not so easy to determine how different the targets should appear if edge integration had a strength of zero—that is, if target lightness were determined solely by local contrast. The problem stems from a lack of specificity in contrast models, despite their mathematical form. There are several ways to estimate the lightness values that would be produced solely by local contrast.

1. The targets in the simultaneous contrast display have local luminance contrasts that are very different. They are in fact just as different as a black target and a white target on the same homogeneous background. Thus, we can plausibly infer that without edge integration the targets would appear black and white, respectively. Given that black and white differ by 7.5 Munsell steps, we can frame our test as follows. With edge integration at zero strength, the targets should appear to differ by 7.5 Munsell steps. With edge integration at a strength of 100%, the targets should appear to differ by 0 Munsell steps. The obtained value, less than 1 Munsell step, is much closer to the 100% prediction than to either the 0% prediction or to the 50% estimate of Reid and Shapley (1988).

2. A second approach to estimating the appearance of the targets without edge integration is to consider how they would appear if the two halves of the display—that is, the gray target on the white background and the gray target on the black background—were presented in the absence of the other so that no integration from one tar-

get to the other is possible. If each half were presented within an otherwise totally dark environment, the gray target on the white background would continue to appear slightly darker than middle gray. But the gray target on the black background would appear white (presumably because it would now be the highest luminance). In short, the targets would appear to differ by four or five Munsell steps. Thus, this method produces a smaller estimate of the difference in target appearance than would be produced in the absence of edge integration, but even so, the difference that is actually obtained (between half a Munsell step and one Munsell step) comes out closest to the predicted difference (zero Munsell steps) with edge integration at full strength.

3. Edge integration between the two targets would be prevented if the border between the white background and the black background were retinally stabilized. In fact, Piantanida and Gilchrist (1994) have conducted just such an experiment. They found that this modification of the display makes the gray targets appear black and white, respectively. In this case, the local contrast ratio for each of the targets and thus, presumably, the local contrast effect, remains unchanged even as the role of edge integration is excluded. This approach to the question yields the same answer as approach (1) above, namely, obtained lightness values in the simultaneous contrast display are relatively close to what would be expected if edge integration were perfect. Piantanida and Gilchrist's experiment was anticipated in analogous experiments conducted earlier by Yarbus (1967) and Arend et al. (1971).

The strength of edge integration was quantified in Gilchrist (1988). Observers viewed a configuration consisting of two adjacent target/surround patterns and selected a luminance value for one target that matched the other target in lightness. By altering whether the boundary between the two surrounds appeared as a reflectance edge or an illuminance edge, the display could be made to appear as a contrast display or as a constancy display. When that boundary appeared as an illuminance edge, the matches made by the observers approximated a ratio match. But when the boundary appeared as a reflectance edge, the matches approximated a luminance match; specifically, the average match was five times closer to a luminance match than to a ratio match. If edge integration were perfect, perfect luminance matches would result. But if lightness perception were completely determined locally, with no contribution by edge integration, perfect ratio matches should result. Thus, by this measure, the strength of edge integration is 5 out of 6, or 83%.

In a somewhat analogous experiment reported by Gilchrist et al. (1983), the results showed edge integration to have a strength of 93%. McCann (1987) reported data that are equivalent to an edge integration strength of 75% by defining the local predictions in terms of perceived values if the two halves of the simultaneous contrast display were each presented in isolation. As pointed out in analysis (2) above, this method gives the smallest estimate of the difference in target appearance predicted

by a local process. Arend and Spehar (1993) conducted an extensive parametric investigation using center/surround patterns embedded in Mondrians, very similar to the kind of patterns Reid and Shapley have used to support their arguments. Arend and Spehar's displays also allow a very straightforward definition of local as well as global predictions. They obtained results that support the strong version of edge integration as opposed to Reid and Shapley's (1988) weaker version.

Our findings present a challenge as well to contrast models based on lateral inhibitory concepts. Those models (Cornsweet, 1970; Heinemann, 1955; Jameson & Hurvich, 1964; Leibowitz et al., 1953) predict a very steep fall-off of contrast effects with retinal displacement between so-called induced and inducing regions.

Leibowitz et al. (1953) did obtain a reduction of the darkening effect on one isolated region by another isolated region as the retinal distance between them was increased. But their stimulus regions were presented against darkness, under conditions known (Katz, 1935; Kaufman, 1974; Rock, 1975) to produce a luminous appearance. We do not necessarily expect our conclusions to apply to such disembodied luminous patches. The "inducing" and "induced" regions in our experiments are seen as parts of a continuous planar surface. This situation, one that is typical of natural, complex images, is just the kind of situation for which edge integration seems ideally suited and to which our results should logically apply.

A central feature of visual perception is the achievement of perceptual constancy in the face of an impressive array of environmental challenges. Perceived properties of the world remain remarkably stable even though changes such as those of viewing distance and level of illumination produce dramatic variations in the pattern of stimulation reaching the receptor surface. In general, perceptual qualities depend upon high order relationships within the configuration, and it matters little what particular part, or how much, of the retina is stimulated. The kind of spatial function implicit in contrast models of lateral inhibition would tie our perception of lightness far too much to the retinal locus of stimulation. This would produce major failures of lightness constancy that do not in fact exist. Claims that the effectiveness of edge integration is reduced with distance imply the same kind of perceptual instability.

The ability of the visual system to integrate edges confers on it a high degree of indifference to chance variations in which surfaces are adjacent to which, either in the distal stimulus or in the proximal stimulus. Every surface in a group of adjacent, coplanar surfaces seems to be effortlessly related, at least potentially, to every other surface.

Our work testing edge integration has turned up two dramatic examples of the rescaling of luminance ratios. In all of our experiments involving the presentation of one to five squares within a spotlight in an otherwise dimly lit room, we found that the range of perceived lightness values (log luminance range, 0.4) is about one

third the range of physical reflectance values (log luminance range, 1.48) in the stimulus. On the other hand, we found in the Mondrian world that a fourfold range of reflectance values was perceived as a 6.5:1 range of lightness values. Both of these outcomes, range compression in the first case and range expansion in the second, represent apparent failures of the ratio principle, defined here as a one-to-one relationship between the stimulus range and the perceived range. There is some parallel between these results and the effect that has come to be called *contrast–contrast*. Chubb, Sperling, and Solomon (1989) showed that the apparent contrast in a random-dot pattern depends very much on the physical contrast of the surrounding region. Our results, though, are much larger in magnitude than those of Chubb et al., and the conditions of our experiments do not seem to satisfy the conditions that produce contrast–contrast. Other than this connection, we are unaware of any systematic treatment of, or even discussion of, such dramatic compression or expansion effects. We believe a systematic study of the conditions under which these rescaling effects occur might reveal important insights regarding the processing of surface lightness, particularly in complex images.

REFERENCES

Arend, L. E. (1973). Spatial differential and integral operations in human vision: Implications of stabilized retinal image fading. *Psychological Review*, **80**, 374-395.

Arend, L. E., Buehler, J. N., & Lockhead, G. R. (1971). Difference information in brightness perception. *Perception & Psychophysics*, **9**, 367-370.

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

Berman, P. W., & Leibowitz, H. W. (1965). Some effects of contour on simultaneous brightness contrast. *Journal of Experimental Psychology*, **69**, 251-256.

Bonato, F., & Gilchrist, A. L. (in press). The perception of luminosity on different backgrounds and in different illuminations. *Perception*.

Bruno, N. (1992). Lightness, equivalent backgrounds and the spatial integration of luminance. *Perception*, **21** (Suppl.), 80.

Burgh, P., & Grindley, G. C. (1962). Size of test patch and simultaneous contrast. *Quarterly Journal of Experimental Psychology*, **14**, 89-93.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences*, **86**, 9631-9635.

Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press.

Evans, R. M. (1948). *An introduction to color*. New York: Wiley.

Gelb, A. (1929). Die "Farbenkonstanz" der Sehdinge. In W. A. von Bethe (Ed.), *Handbuch normalen und pathologischen Psychologie* (pp. 594-678). Berlin: Springer-Verlag.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics*, **28**, 527-538.

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

Gilchrist, A. L., & Bonato, F. (in press). Anchoring of lightness values in center/surround displays. *Journal of Experimental Psychology: Human Perception & Performance*.

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

Gilchrist, A. L., & Jacobsen, A. (1983). Lightness constancy through a veiling luminance. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 936-944.

Gilchrist, A. [L.], & Jacobsen, A. (1984). Perception of lightness and illumination in a world of one reflectance. *Perception*, **13**, 5-19.

Gogel, W. C., & Mershon, D. H. (1969). Depth adjacency in simultaneous contrast. *Perception & Psychophysics*, **5**, 13-17.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing- and test-field luminances. *Journal of Experimental Psychology*, **50**, 89-96.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Horn, B. K. P. (1986). *Robot vision*. Cambridge, MA: MIT Press.

Jameson, D., & Hurvich, L. M. (1964). Theory of brightness and color contrast in human vision. *Vision Research*, **4**, 135-154.

Katz, D. (1935). *The world of colour*. London: Kegan Paul, Trench, Trubner & Co.

Kaufman, L. (1974). *Sight and mind: An introduction to visual perception*. New York: Oxford University Press.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, **61**, 1-11.

Leibowitz, H., Mote, F. A., & Thurlow, W. R. (1953). Simultaneous contrast as a function of separation between test and inducing fields. *Journal of Experimental Psychology*, **46**, 453-456.

McCann, J. J. (1987). Local/global mechanisms for color constancy. *Die Farbe*, **34**, 275-283.

McCann, J. J., & Savoy, R. L. (1991). Measurements of lightness: Dependence on the position of a white in the field of view. *SPIE*, **1453**, 402-411.

Mikesell, W. H., & Bentley, M. (1930). Configuration and contrast. *Journal of Experimental Psychology*, **13**, 1-23.

Newson, L. J. (1958). Some principles governing changes in the apparent lightness of test surfaces isolated from their normal backgrounds. *Quarterly Journal of Experimental Psychology*, **10**, 82-95.

Piantanida, T., & Gilchrist, A. (1994). *Direct evidence for edge integration in lightness perception*. Manuscript in preparation.

Reid, R. C., & Shapley, R. (1988). Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Research*, **28**, 115-132.

Rock, I. (1975). *An introduction to perception*. New York: Macmillan.

Shapley, R. (1985). The importance of contrast in the perception of brightness and form. *Transactions of the American Philosophical Society*, **75**, 20-29.

Shapley, R., & Reid, R. C. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of Sciences*, **82**, 5983-5986.

Stewart, E. (1959). The Gelb effect. *Journal of Experimental Psychology*, **57**, 235-242.

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

Wallach, H. (1963). The perception of neutral colors. *Scientific American*, **208**, 107-116.

Wallach, H. (1976). *On perception*. New York: Quadrangle/The New York Times Book Co.

Whittle, P., & Challands, P. D. C. (1969). The effect of background luminance on the brightness of flashes. *Vision Research*, **9**, 1095-1110.

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.