

by Edwin H. Land

VOL 237 NO 6 P 108-128



PUBLISHED BY W. H. FREEMAN AND COMPANY 660 MARKET STREET, SAN FRANCISCO, CALIFORNIA 94104

Copyright ©1977 by Scientific American Inc. All rights reserved. Printed in the U. S. A No part of this offprint may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher.

The Retinex Theory of Color Vision

A retina-and-cortex system (retinex) may treat a color as a code for a three-part report from the retina, independent of the flux of radiant energy but correlated with the reflectance of objects

by Edwin H. Land

he scientific tradition of simplifying the conditions of an experiment has left us until recently

without a satisfactory explanation of how the eye sees color in everyday life. Paradoxically the modern technology of color photography has reinforced the belief that the colors discerned by Newton in the spectrum are, with minor qualifications, the colors of the world around us. We know, for example, that if we use daylight color film when we take a picture in the light shed by an ordinary tungsten-filament lamp, the picture will turn out to have a strong reddish cast. That, we say, is because the rays from the tungsten filament are too "red." never asking how we ourselves can move constantly in and out of tungsten-lit worlds without experiencing any change in the color of familiar objects: apples, lemons, strawberries, bread, human faces (the tones of which are so hard to get right on a television screen).

How, then, does the eye deal with the excess of "red" in a tungsten-lit room? As I hope to demonstrate in this article, the eye, in determining color, never perceives the extra red because it does not depend on the flux of radiant energy reaching it. The eye has evolved to see the world in unchanging colors, regardless of always unpredictable, shifting and uneven illumination. How the eye achieves this remarkable feat has fascinated me for many years.

In 1959 I described in these pages a series of experiments in which a scene created by the superposition of two black-and-white transparencies, one projected through a red filter and the other projected without a filter (that is, in white light), conveys to the eye nearly the gamut of colors present in the original scene [see "Experiments in Color Vision." by Edwin H. Land: SCIENTIFIC AMERICAN Offprint No. 223]. To produce such "red-and-white" images the picture projected through the red filter is

taken through a red filter and the picture projected in white light is taken through a green filter. It would be expected that the superposed image on the projection screen could generate only red, white and various shades of pink. Actually one sees a picture remarkably similar to the full-color photograph reproduced on the opposite page. In the red-and-white photographic projection peppers are green, radishes and strawberries are red, the orange is orange, the lemon and bananas are pale yellow, the wood cutting board and knife handle are brown and the design on the plate is blue.

The challenge presented by our early red-and-white experiments led us step by step over a 20-year period to an explanation of how the visual system is able to extract reliable color information from the world around us, a world in which virtually every scene is lighted unevenly, in which the spectral composition of the radiation falling on a scene can vary enormously and in which illumination as brief as a lightning flash suffices for the accurate identification of color. If the nature of the responses of the photoreceptors in the retina of the eye even approximated what most of us were taught in school, functioning primarily as intensity-level meters with peaks in three different parts of the spectrum, we would be continually confusing one color with another. An object that looked yellow in one part of our field of view might look green or gray or even red when moved to a different part of the field. The fact remains that objects retain their color identity under a great variety of lighting conditions. This constancy is not a minor second-order effect but is so fundamental as to call for a new description of how we see color.

he visual pigments are photosensitive molecules that respond to a wide band of light frequencies. The three pigments in the cone cells of the

retina cover the visible spectrum in three broad, overlapping curves. The pigment with a peak sensitivity at a wavelength of 440 nanometers responds in some degree to the entire lower-frequency half of the visible spectrum. Each of the other two pigments responds to almost two-thirds of the visible spectrum, the two being offset at their peaks by barely 30 nanometers, with their peak sensitivities located at 535 and 565 nanometers [see upper illustration on page 4].

In this discussion the names of colors—"red", "green", "blue" and so on—will be reserved for the color sensation we have when we look at the world around us. In short, only our eyes can categorize the color of objects; spectrophotometers cannot. This point is not a trivial one because many people viewing some of our experiments for the first time will identify something as being red or green but will then ask, as if their eyes were being fooled, "What color is it really?" The answer is that the eye is not being fooled. It is functioning exactly as it must with involuntary reliability to see constant colors in a world illuminated by shifting and unpredictable fluxes of radiant energy.

Since I believe the study of color in fully colored images is best begun by examining images that are completely devoid of and completely uncomplicated by the experience of color, let me describe that experience in some detail. The hypersensitive system based on the rod cells in the retina functions at light levels as much as 1,000 times weaker than the systems based on the cone cells do, so that it is possible to answer the interesting question: What colors will one see if only the rod system is activated? One procedure is to put on a pair of tightly fitting goggles equipped with neutral-density filters that reduce the incident light by a factor of 30,000. After one has worn the goggles for about half

can
re
ook
that
-er
ects
f d a
rad



STILL LIFE was used to make the four black-and-white images presented below. The reproduction of the still life above was made by conventional processes of color photography and photoengraving to show the reader what the colors of the original objects in the scene

were. The black-and-white images were made with film-filter combinations that closely duplicate the separate wavelength sensitivities of the four systems of photoreceptors in the retina of the eye: the three systems of cone cells and the hypersensitive system of rod cells.



BLACK-AND-WHITE IMAGES OF STILL LIFE were taken with four different film-filter combinations, creating what the author calls retinex records. The picture at the top left was taken with a film whose spectral response was altered so that it matched the spectral sensitivity of the long-wave cone pigments in the eye. This photograph enables the observer to see a colorless image that approximates the image produced by the long-wave cones by themselves. The pic-

ture at the top right shows the same scene as it would be viewed by the middle-wave cone pigment. The picture at the bottom left is the scene as it would be viewed by the short-wave cone pigment. The picture at bottom right corresponds to the image seen by the rods. Unlike cone images, which cannot be viewed independently, images produced by the rod pigment can be studied in isolation at very low light levels, without interference from much less sensitive cone systems.

that each of the objects, the pieces of paper for example, whether dark or light or in between, maintains its lightness without significant change as it is moved around the room into regions of higher or lower flux. Light papers will be seen as being light and dark papers simultaneously as being dark, even with the same flux coming from each of them to the eye. Strong gradients of flux across the field will be apparent only weakly, if at all.

Furthermore, in an intricate collage of areas of various lightnesses sizes and shapes, the lightness of a given element does not change visibly as it is relocated

in any part of the collage and associated with a new arbitrary surround. When a small area is totally surrounded by a large area, the lightness of the small area will change somewhat depending on whether the large area is darker or lighter than the small one. In general, however, the impressive fact is that the lightness of a given area is not appreciably modified by the immediately surrounding areas, nor is it modified by the still larger areas surrounding them.

Although I have been describing a colorless world as it is seen by the hypersensitive receptors of rod vision, all

the observations about the stability of lightness values can readily be reproduced with a montage of white, black and gray papers viewed at ordinary light levels. If, for example, a square of matte-surface black paper or, better still, black velvet is placed at one side of such a montage and a square of white paper is placed at the opposite side several feet away, with an assortment of light and dark papers scattered in between, one can place a strong light source close enough to the black square so that it sends more radiant energy to the eye than the white square, remote from the light: yet the black square will continue to look black and the white square white. In fact, with the montage still strongly illuminated from one side either the black square or the white one can be moved to any other part of the montage without a significant change in its appearance.

This remarkable ability of the eye to discover lightness values independent of flux, so convincingly demonstrated when only a single photoreceptor system is operating, is the rock on which a satisfactory description of color vision can be built. The first response of the visual system is for the receptors to absorb the light falling on the retina. Whereas the initial signal produced in the outer segment of the receptor cell is apparently proportional to the light flux absorbed by the visual pigment, the final comprehensive response of the visual

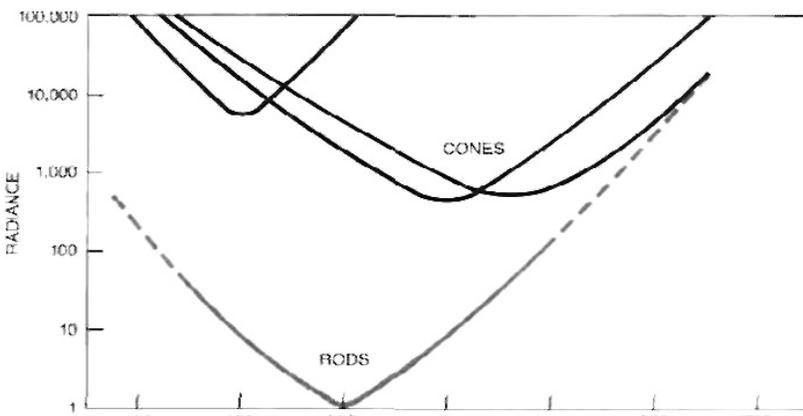
system is "lightness," which shows little or no relation to the light flux absorbed by the visual pigment.

The processing of fluxes to generate lightnesses could occur in the retina, or in the cerebral cortex, or partially in both. Since we are uncertain of the location of the mechanisms that mediate these processes, I have coined the term *retinex* (a combination of retina and cortex) to describe the ensemble of biological mechanisms that convert flux into a pattern of lightnesses. I shall therefore use the term throughout this article in referring to these biological mechanisms. I shall also reserve the term *lightness* to mean the sensation produced by a biological system. Although the rods can be stimulated at light intensities below the cone threshold, the cones cannot be stimulated without exciting the rods. For cones we must study the lightness images produced by each individual set of receptors using *retinex* photography, as I shall explain below, or learn the properties of lightness images from model calculations based on spectroradiometric measurements.

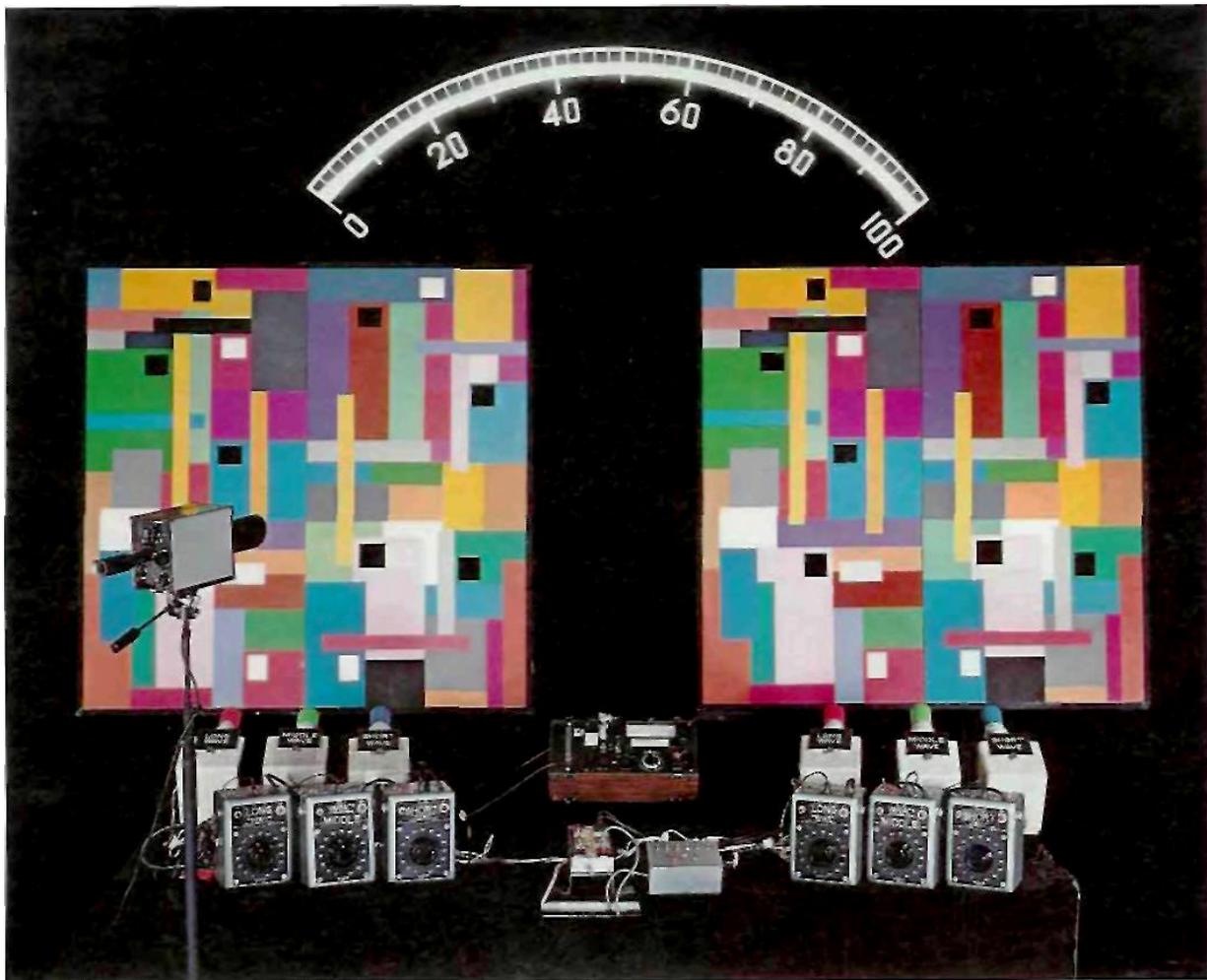
Now that we know that at low light levels an isolated receptor system generates an image in terms of lightness that is completely free of color, might it be possible to bring one of the cone systems into operation along with the hypersensitive system, so that only the

WAVELENGTH (NANOMETERS)

NORMALIZED SPECTRAL SENSITIVITIES OF FOUR VISUAL PIGMENTS (*solid lines*) span the visual spectrum in overlapping curves. Curve that peaks at about 500 nanometers corresponds to sensitivity of rod pigment. Other three curves represent cone pigments. Broken lines show sensitivities of the film-filter combinations that were selected to match the sensitivities of the four retinal pigments and used to make the black-and-white retinex records in the illustration at bottom of preceding page. Cone curves are adapted from work of Paul Brown and George Wald of Harvard University. The rod curve is standard scotopic luminosity curve.



sensitive rod system provides vision at radiance levels about 1000 times weaker than the light levels needed to activate the cone systems. It has been shown in author's laboratory that multi-colored scenes exhibit nearly their normal range of colors when they are viewed at light levels so adjusted that only rod system and one cone system, the long-wave system, are responding.



"COLOR MONDRIAN" EXPERIMENT employs two identical displays of sheets of colored paper mounted on boards four and a half feet square. The colored papers have a matte finish to minimize specular reflection. Each "Mondrian" is illuminated with its own set of three projector illuminators equipped with band-pass filters and independent brightness controls so that the long-wave ("red"), middle-wave ("green") and short-wave ("blue") illumination can be mixed in any desired ratio. A telescopic photometer can be pointed at any area to measure the flux, one wave band at a time, coming to the eye

from that area. The photometer reading is projected onto the scale above the two displays. In a typical experiment the illuminators can be adjusted so that the white area in the Mondrian at the left and the green area (or some other area) in the Mondrian at the right are both sending the same triplet of radiant energies to the eye. The actual radiant-energy fluxes cannot be re-created here because of the limitations of color reproduction. Under actual viewing conditions white area continues to look white and green area continues to look green even though the eye is receiving the same flux triplet from both areas.

LONG WAVE 5.8
 MIDDLE WAVE 3.2
 SHORT WAVE 1.6 ENERGY
 AT EYE
 (MILLIWATTS PER
 STERADIAN PER SQUARE METER)

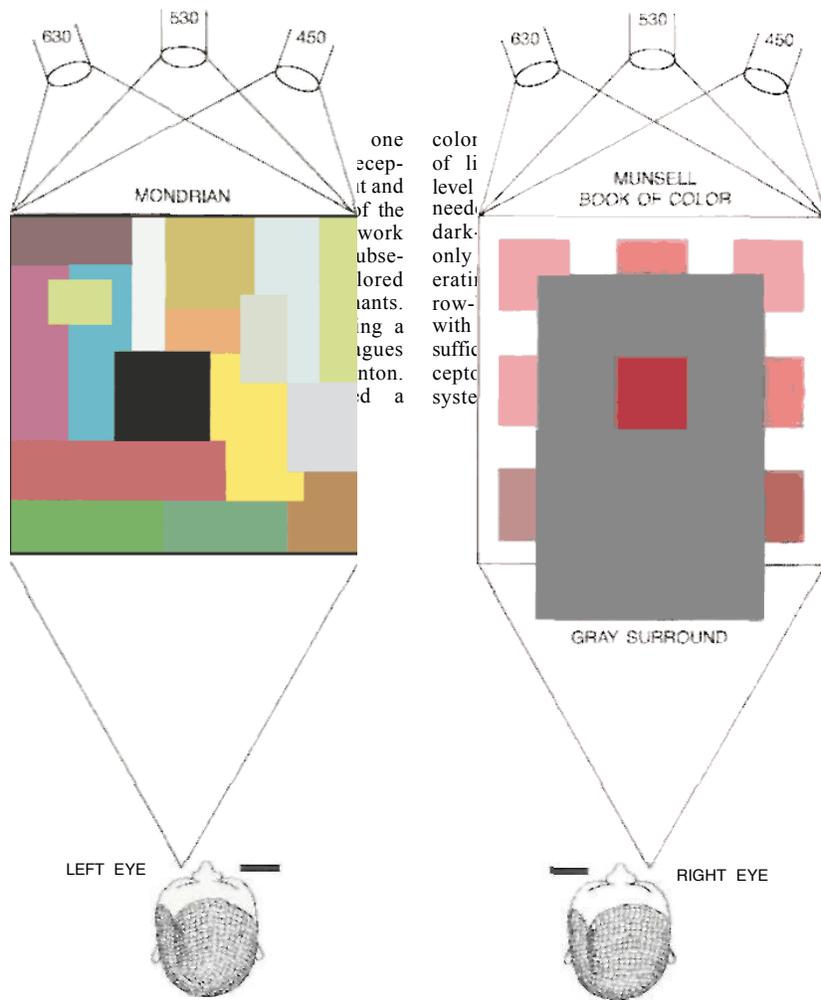
LONG WAVE 5.8
 MIDDLE WAVE 3.2
 SHORT WAVE 1.6 ENERGY
 AT EYE
 (MILLIWATTS PER
 STERADIAN PER SQUARE METER)

LONG WAVE 5.8
 MIDDLE WAVE 3.2
 SHORT WAVE 1.6
 ENERGY AT EYE
 (MILLIWATTS PER
 STERADIAN PER SQUARE METER)

IDENTICAL ENERGY FLUXES AT THE EYE provide different color sensations in the Mondrian experiments. In this example, with the illuminators from the long-wave, middle-wave and short-wave il-

luminators adjusted as indicated, an area that looks red continues to look red (left), an area that looks blue continues to look blue (middle)

and an area that looks green continues to look green (*right*), even though all three are sending to the eye the same triplet of long-, middle- and short-wave energies. The same triplet can be made to come from any other area: if the area is white, it remains white; if the area is gray, it remains gray; if it is yellow, it remains yellow, and so on.



COLOR-MATCHING EXPERIMENT uses a simplified Mondrian of 17 color areas (left) and a standard color reference, *The Munsell Book of Color*, which contains 1150 color "chips" (right). The Mondrian is illuminated with three narrow-band light sources: one at 630 nanometers (long-wave light), one at 530 nanometers (middle-wave light) and one at 450 nanometers (short-wave light). The ratio of the three illuminants can be adjusted so that the triplet of energies reflected to the eye from any chosen area will exactly equal the triplet that previously reached the eye from some other area. In this experiment five areas, gray, red, yellow, blue and green, were selected in sequence to send the same triplet of energies to the eye. In each of the five consecutive parts of this experiment the observer selected from the Munsell book the chips that came closest to matching the 17 areas of the Mondrian. The Munsell book was illuminated throughout the experiment with a constant spectral mixture of three narrow-band lights adjusted at the outset so that the white Munsell chip appeared the "best white." The experiment was set up so that the observers used one eye for viewing the Mondrian and the other eye for viewing chips. Gray paper with an opening was used to provide chips with a constant surround.

receptor systems, namely the rods and the long-wave cones, were receiving enough light to function.

The resulting image exhibited a remarkable range of color, enabling an observer to assign to each area in the display the same color name it would have if it were illuminated above the cone threshold. The result is reminiscent of the multicolored images produced by the red-and-white system. The demonstration explicitly confirms our early proposition that the lightness information collected at two wave bands by separate receptor systems is not averaged, point by point and area by area, but is kept distinct and is compared. We know that the rod system does not produce a colored image when the image is seen by itself, and we know that the long-wave light alone cannot produce an image with a variety of colors. The combination, however, gives rise to a wide variety of colors, namely reds, yellows, browns, blue-greens, grays and blacks.

What, then, accounts for the color? The emergence of variegated colors can be ascribed to a process operating somewhere along the visual pathway that compares the lightnesses of the separate images on two wave bands, provided by the two independent retinex systems. The two-receptor experiment makes it plausible that when three independent images constituting the lightnesses of the short-, middle- and long-wave sets of receptors are associated to give a full-colored image, it is the comparison of the respective lightnesses, region by region, that determines the color of each region. The reason the color at any point in an image is essentially independent of the ratio of the three fluxes on three wave bands is that color depends only on the lightness in each wave band and lightness is independent of flux.

As we have seen, the spectral sensitivities of the visual pigments overlap broadly. If we illuminated a scene with the entire range of wavelengths to which a single visual pigment is sensitive, we would see a large variety of colors because more than one retinex system would respond. With the help of filters and appropriate film emulsions, however, we can isolate the lightnesses that would ordinarily be incorporated into the sensation of color. We call black-and-white photographs made for this purpose retinex records.

The photographic technique, making use of silver emulsions, performs two functions. First, the system provides spectral sensitivities that are the same as those of the visual pigments. Second, it generates black-and-white pictures for a human observer to examine. It is the human visual system that converts the photographic pattern deposited in silver into lightness. Ideally we should like

ILLUMINANT
(MILLIWATTS PER
STERADIAN PER

WAVELENGTHS
(NANOMETERS)

WAVELENGTHS
(NANOMETERS)

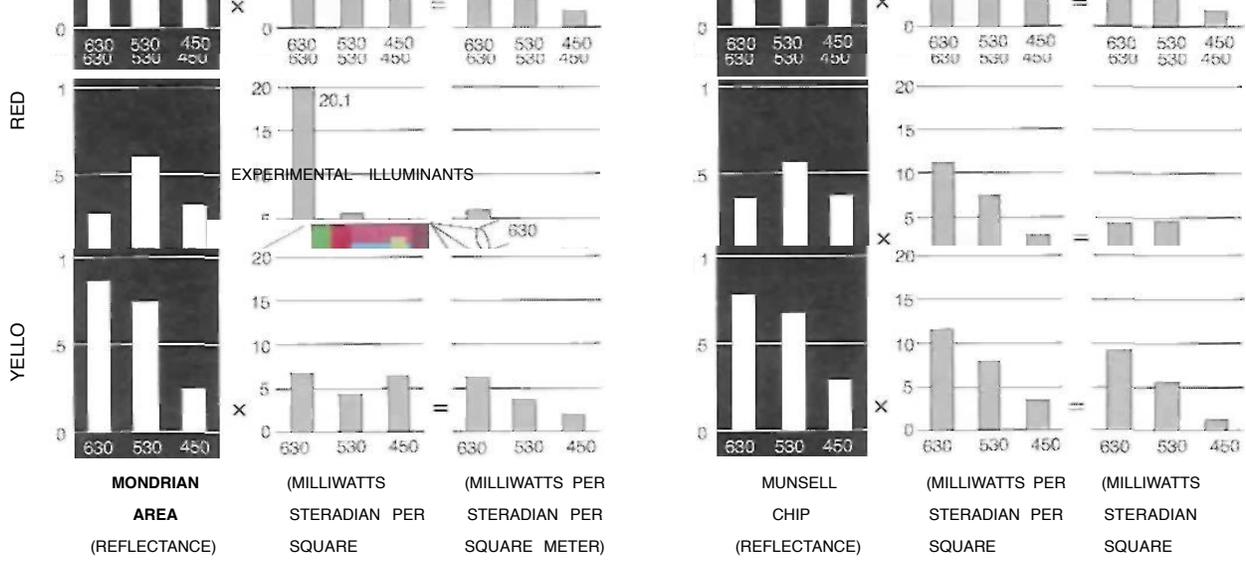
WAVELENGTHS
(NANOMETERS)

WAVELENGTHS
(NANOMETERS)

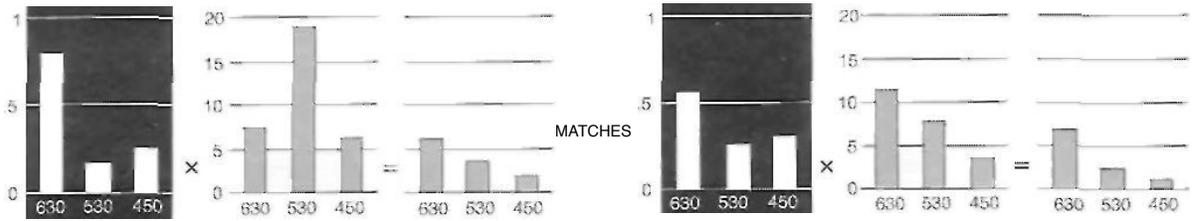
WAVELENGTHS
(NANOMETERS)

MUNSELL CHIPS SELECTED BY OBSERVERS to match the five Mondrian areas that had sent identical triplets of energy to the eye are reproduced. The Munsell book was illuminated with a constant spectral mixture of narrow-band illuminants (*bars at top*) and the chips were viewed within a constant gray surround. The energy that

was sent to the eye by the selected Munsell chips is shown by the bars at the bottom of the illustration. It is evident that the match between the Mondrian areas and the Munsell chips is not made on the basis of the flux of radiant energy at the eye of the observer. What does cause the two areas to match is described in the illustrations that follow.



MATCHES



MATCHES

MATCHES

MATCHES

(NANOMETERS)

(NANOMETERS)

(NANOMETERS)

(NANOMETERS)

WAVELENGTHS
(NANOMETERS)

(NANOMETER)

FURTHER ANALYSIS OF MATCHING EXPERIMENT begins to identify the basis on which the visual system makes the color match between the Mondrian area and the Munsell chip without regard to the

flux each member of the pair sends to the eye. The efficiency with which a given area in the Mondrian reflects light in each of the three wave bands (first column) multiplied by the amount of energy striking that area in each of the wave bands (second column)

yields

the energy triplet that reaches the eye (*third column*). The three columns at the right contain comparable data for the Munsell chips selected as a match for the Mondrian areas. Whereas illustration at bottom of the preceding page shows that the eye does not match colors using a "meter" that measures triplets of energies at the eye, this illustration shows that when a match is made, it is the reflectances of two areas that correspond, as is shown in first and fourth columns.

8

our observer to examine the black-and-white pattern with only one set of cones. reporting the lightnesses appropriate to that set. At any point in the black-and-white pattern, however, the reflectance is essentially the same throughout the visible spectrum. Therefore with a black-and-white photograph we stimulate all the receptors with the same information. that is. with the energies that would be absorbed by a single visual pigment. If we assume that all the retinex systems process information in an identical manner, we can propose that sending this identical information to several sets of receptors is the same as sending it to only one receptor, thereby enabling us to see what the image would look like if it were possible to isolate it.

On page 3 the reader will see three black-and-white pictures taken through retinex filters that simulate the response of the three cone pigments. The strawberries and radishes, for example, are light on the long-wave record, darker on the middle-wave record and darkest on the short-wave record. Although the orange and lemon are about as dark as the strawberries and radishes on the short-wave record, they are nearly as light on the middle-wave record as they are on the long-wave record. On the printed page the distinctions are subtle. To the eye viewing an actual full-color scene the subtle distinctions provide all the information needed to distinguish countless shades and tints of every color.

After the three lightnesses of an area have been determined by the three retinex systems no further information is necessary to characterize the color of any object in the field of view. Any specific color is a report on a trio of three specific lightnesses. For each trio of lightnesses there is a specific and unique color.

The limitations of color photography make it impossible to show the reader the demonstrations readily accomplished in our laboratory, which dramatically reveal the independence of perceived color from the flux reaching the eye. What the reader would see would be two boards four and a half feet square identically covered with about 100 pieces of paper of various colors and shapes. In order to minimize the role of specular reflectance the papers have matte surfaces and, except for black, have a minimum reflectance of at least 10 percent for any part of the visible spectrum. In these displays, which we call "color Mondrians" (after the Dutch painter to whose work they bear a certain resemblance), the papers are arranged so that each one is surrounded by at least five or six others of different colors [see top illustration on page 5].

Each of the identical Mondrians is illuminated by its own set of three pro-

jectors equipped with sharply cutting band-pass filters (not retinex filters): one at 670 nanometers embracing a band of long waves, one at 540 nanometers embracing a band of middle waves and one at 450 nanometers embracing a band of short waves. The amount of light from each illuminating projector is controlled by a separate variable transformer. In addition the illuminating projectors have synchronized solenoid-activated shutters to control the duration of illumination. There is a telescopic photometer that can be precisely aimed at any region of either Mondrian to measure the amount of radiation reflected from any point and therefore the amount of flux reaching the eye. The output of the photometer is projected on a scale above the Mondrian, where it can be seen by those taking part in the demonstration.

The demonstration begins with the three illuminating projectors turned on the Mondrian on the left; the Mondrian on the right remains dark. The variable transformers are set so that the entire array of papers in the left Mondrian are deeply colored and at the same time the whites are good whites. This setting is not critical. Then, using one projector at a time and hence only one wave band at a time, we measure with the telescopic photometer the energy reaching the eye from some particular area, say a white rectangle. The readings from the white area (in milliwatts per steradian per square meter) are 65 units of long-wave light, 30 units of middle-wave light and five units of short-wave light. We have now established the three energies associated with that sensation of white.

We turn off the three projectors illuminating the color Mondrian on the left. On the right we turn on only the long-wave projector. We select a different area of unknown color and adjust the long-wave light until the long-wave energy coming to the eye from the selected area is the same as the long-wave energy that a moment ago came from the white paper in the Mondrian on the left, 65 units. We turn off the long-wave projector and separately adjust the transformers controlling the middle- and short-wave projectors, one after the other, so that the energies sent to the eye from the selected area are also the same as those that came from the white area on the left. We have not yet turned on all three light sources simultaneously, but we know that when we do so, the triplet of energies reaching the eye from the selected area of still unknown color will be identical with the triplet that had previously produced the sensation white.

When we turn on the three illuminants, we discover that the area in the Mondrian on the right is green. We now illuminate the Mondrian on the left with its illuminants at their original settings

so that both Mondrians can be viewed simultaneously. The white area on the left continues to look white and the green area on the right continues to look green, yet both are sending to the eye the same triplet of energies: 65, 30 and five in the chosen units.

We turn off the illuminants for both Mondrians and select some other area in the left Mondrian and sequentially adjust the energies reaching the eye from it so that they are the same as the energies that originally gave rise to the sensation of white and also gave rise to the sensation of green in the right Mondrian. When we turn on all three projectors illuminating the left Mondrian, we see that this time the selected area is yellow. The triplet of energies reaching our eye is the same one that had previously produced the sensations of white and green. Again, if we wish, the yellow and green can be viewed simultaneously, with yellow on the left and green on the right.

We can continue the demonstration with other areas such as blue, gray, red and so on. It is dramatically demonstrated that the sensation of color is not related to the product of reflectance times illumination, namely energy, although that product appears to be the only information reaching the eye from the various areas in the Mondrians.

In order to demonstrate that the color sensations in these experiments do not involve extensive chromatic adaptation of retinal pigments the projectors are equipped with synchronized shutters so that the Mondrians can be viewed in a brief flash, a tenth of a second or less in duration. Regardless of the brevity of observation the results of the demonstrations are not altered. Thus one can say that neither chromatic adaptation nor eye motion is involved in producing the observed colors. Finally, the very essence of the design of the color Mondrian is to obviate the significance of the shape and size of surrounding areas, of the familiarity of objects and of the memory of color. Curiously, from time to time there is a casual attempt to adduce what is called color constancy as an explanation of these demonstrations.

Clearly color constancy is only a compact designation of the remarkable competence that is the subject of this article.

The mystery is how we can all agree with precision on the colors we see when there is no obvious physical quantity at a point that will enable us to specify the color of an object. Indeed, one can say the stimulus for the color of a point in an area is not the radiation from that point. The task of psychophysics is to find the nature of the stimulus for that color.

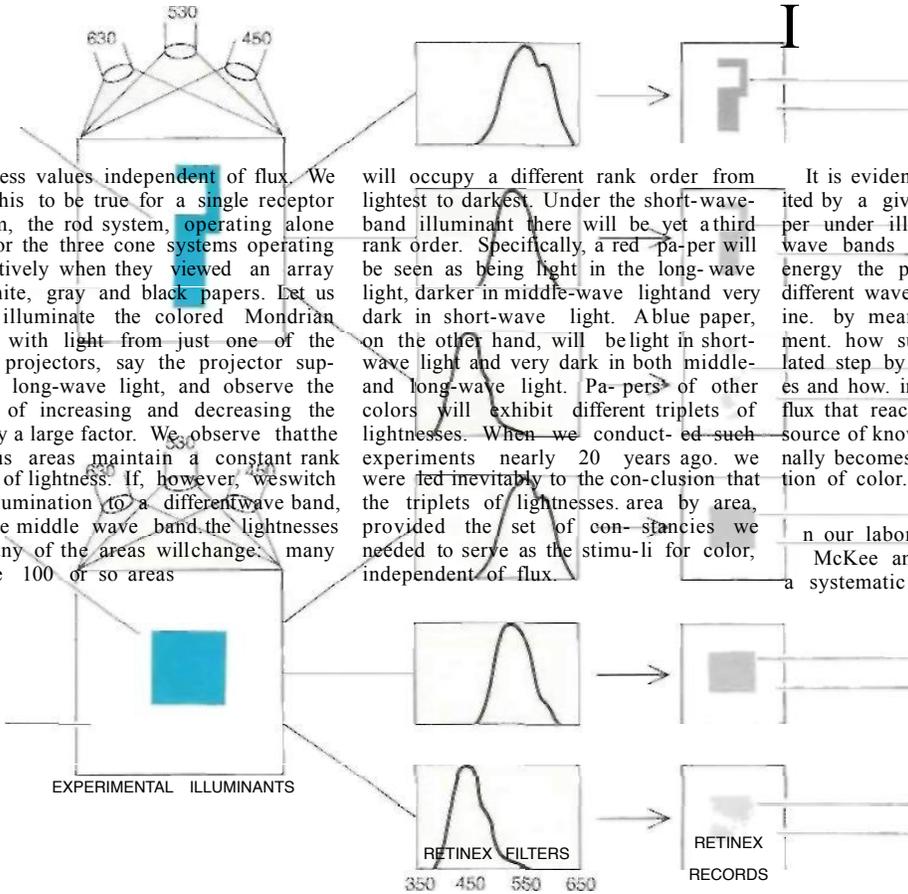
Here let us remember that what the eye does unfailingly well is to discover

lightness values independent of flux. We saw this to be true for a single receptor system, the rod system, operating alone and for the three cone systems operating collectively when they viewed an array of white, gray and black papers. Let us now illuminate the colored Mondrian array with light from just one of the three projectors, say the projector supplying long-wave light, and observe the effect of increasing and decreasing the flux by a large factor. We observe that the various areas maintain a constant rank order of lightness. If, however, we switch the illumination to a different wave band, say the middle wave band, the lightnesses of many of the areas will change; many of the 100 or so areas

will occupy a different rank order from lightest to darkest. Under the short-wave-band illuminant there will be yet a third rank order. Specifically, a red paper will be seen as being light in the long-wave light, darker in middle-wave light and very dark in short-wave light. A blue paper, on the other hand, will be light in short-wave light and very dark in both middle- and long-wave light. Papers of other colors will exhibit different triplets of lightnesses. When we conducted such experiments nearly 20 years ago, we were led inevitably to the conclusion that the triplets of lightnesses, area by area, provided the set of constancies we needed to serve as the stimuli for color, independent of flux.

It is evident that the lightnesses exhibited by a given piece of colored paper under illuminants of three different wave bands is related to the amount of energy the paper reflects to the eye at different wavelengths. Let us now examine, by means of a particular experiment, how such reflectances can be related step by step to perceived lightnesses and how, in the process, the radiant flux that reaches the eye — the ultimate source of knowledge about lightness — finally becomes irrelevant to the sensation of color.

In our laboratory McCann, Suzanne P. McKee and Thomas H. Taylor made a systematic study of observers' re-



ILLUMINANT	WAVE BAND	RATIO	INTEG	RADIA	INTEG	REFL	SCALI	INTEG	REFL
BLUE AREA IN MONDRIAN	LONG WAVE	27.3	100		27.3			5.8	
	MIDDLE WAVE	35.9	100		35.9			6.5	
WHITE PAPER	MIDDLE WAVE	60.7	100		60.7			8.1	
	SHORT WAVE	34.6	100		34.6			6.4	
STANDARD "WHITE" ILLUMINANT	LONG WAVE	38.5	100					38.5	
	MIDDLE WAVE	57.1	100					57.1	

SHORT WAVE

WAVELENGTH
(NANOMETERS)

ROLE OF REFLECTANCE and its psychophysical correlate, lightness, in guiding the eye to match Munsell chips with Mondrian areas was examined with the help of retinex filter-photomultiplier combinations that match the spectral sensitivity of the cone pigments. Under each combination of illuminants (top) the integrated radiance, or flux, in each retinex wave band of a Mondrian area was compared with the integrated radiance of a sheet of white paper. The ratio of

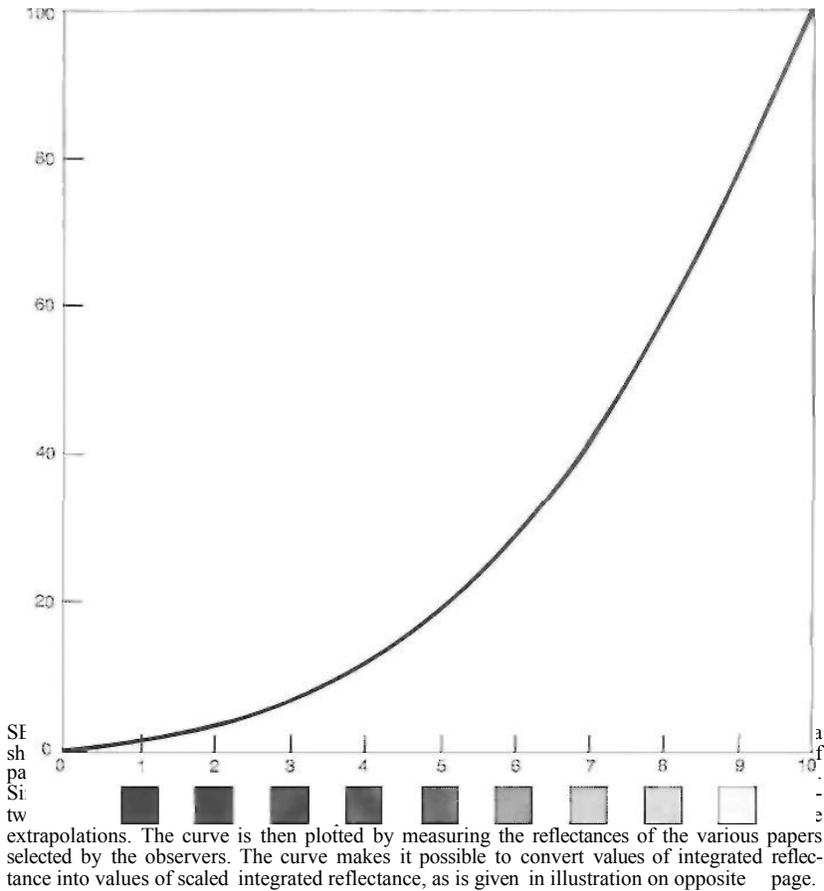
integrated radiances yields the integrated reflectance of the Mondrian area, expressed here in percent. For the matching Munsell chip a set of ratios was similarly determined (*bottom*). The final step in deriving a physical equivalent of lightness is the scaling, or spacing, of integrated reflectances to be consistent with the spacing of lightness sensations. This transformation is explained in the illustration on the opposite page. The scaled values appear in the column at the right

sponses to a simplified color Mondrian with areas of 17 different colors. They asked the observers to match the 17 areas one at a time under different illuminants with colored squares of paper that had been selected from a standard color-reference book. *The Munsell Book of Color* and that were viewed under a constant "white" illumination.

The illuminants on the Mondrian were adjusted in five separate matching experiments so that five different areas (gray, red, yellow, blue and green) sent to the eye an identical triplet of radiances. The observer began by selecting a matching Munsell "chip" for each of the 17 areas in the Mondrian when the gray area in the Mondrian sent a particular triplet of energies to the eye. Another set of 17 matching Munsell chips was selected when the same triplet was later sent to the eye by a red area in the Mondrian, and the same was done for yellow, blue and green areas under illuminants that supplied the same triplet of energies.

The illustrations on page 7 show the details of the experiment and the five different Munsell colors the observers selected to match the five areas when each area sent to the eye precisely the same triplet of energies. In spite of the constancy of the energy reaching one eye from the Mondrian, each observer, using the other eye, selected Munsell chips that were gray, red, yellow, blue and green.

The constant illumination used in viewing the Munsell book was a triplet of illuminants at three wavelengths that observers judged to produce the "best" white. The actual triplet of wavelengths reaching the eye from the whitest paper in the Munsell book was 11.5 units of long-wave light, 7.8 units of middle-wave light and 3.3 units of short-wave



extrapolations. The curve is then plotted by measuring the reflectances of the various papers selected by the observers. The curve makes it possible to convert values of integrated reflectance into values of scaled integrated reflectance, as is given in illustration on opposite page.

light. The illuminants supplied energy in narrow bands with peaks at 630 nanometers, 530 nanometers and 450 nanometers. A similar triplet of narrow-band illuminants were mixed in various proportions to illuminate the Mondrian.

At this point the reader might ask: Would not a single gray area exhibit a pronounced change in color if the surrounding papers had reflected light of widely differing spectral composition? Could these changes in color account for the results of the Mondrian experiments? The answer to the questions is that no manipulation of surrounding papers in the Mondrian is capable of making the gray paper match the red, yellow, blue and green Munsell papers selected by the observers in the Mondrian experiment.

McCann, John A. Hall and I have examined the matter further by repeating the Mondrian-Munsell experiment in various ways so that the average spectral composition of the light reaching

the eye from the Mondrian and its surround remains the same regardless of the spectral composition of the light needed to establish a constant triplet from area to area. We have done this in one case by surrounding the entire Mondrian with brightly colored papers selected in such a way that they exactly offset the average mixture of wave bands from the Mondrian itself and, more dramatically, by cutting the 17 areas of the Mondrian apart and placing them well separated on the backgrounds of offsetting color. Neither arrangement has any significant effect on the Munsell chips chosen to match the various areas of the Mondrian.

Let us return, then, to the search for the stimulus that guides us so accurately to the correct identification of colors. If it is not a flux of radiant energy at the eye from each point in the field of view, what are the physical correlates of the lightnesses of objects on three separate

wave bands, corresponding to the spectral sensitivities of the cone pigments? Can such a precise physical correlate of lightness be demonstrated?

McCann, McKee and Taylor next measured the radiance, or energy at the eye, of the various Mondrian areas and of the matching Munsell chips by using a photomultiplier in conjunction with a version of the retinex filters. Since the retinex-photomultiplier combination integrates the flux of radiant energy over a broad band of wavelengths, the instrument provides a value we call integrated radiance. McCann and his colleagues then obtained the integrated radiances from a large sheet of white paper placed under each of the experimental illuminants that had been used to light the Mondrian in the chip-matching experiments. If the integrated radiance from a Mondrian area is used as the numerator in a fraction and the integrated radiance from the white paper is used as the denominator, one obtains a value for in-

egrated reflectance, which can be expressed as a percent.

The integrated reflectances for the various Munsell chips are determined in the same manner under the constant "white" illumination. This amounts to measuring the percentage of reflectance using detectors with the same spectral sensitivity as the visual pigments. The results show that the Munsell chip chosen by the eye to match a given Mondrian area will have approximately the same three integrated reflectances as the area. For example, the blue area in the Mondrian has a triplet of integrated reflectances (long-, middle- and short-wave) of 27.3, 35.9 and 60.7 percent. The comparable values for the matched Munsell chip are 34.6, 38.5 and 57.1 percent [see illustration on page 10].

Finally, the integrated reflectances are "scaled" so that their equal spacing is consistent with the equal spacing of lightness sensations. The curve for this transformation is shown in the illustration on the preceding page. Using this curve, we see that the blue area in the Mondrian has a triplet of scaled inte-

egrated reflectances of 5.8, 6.5 and 8.1, whereas the corresponding values for the matching Munsell chip are 6.4, 6.7 and 7.9. If we study the five areas that successively sent identical triplets of energies to the eye and compare their scaled integrated reflectances with those of their matching Munsell chips, we find that all the values are in excellent agreement. In other words, in the triplets of integrated reflectances we have identified a highly accurate physical correlate of color sensations. The data fall along the 45-degree line that describes the locus of perfect correlation [see illustration below].

We have sought a physical correlate for lightness, and we have found that the scaled integrated reflectances of the five areas that sent identical triplets of fluxes to our eyes are the same as those of the matching Munsell chip. This correlation enables us to use scaled integrated reflectances as a measured lightness equivalent. The problem now shifts to one of how the eye derives the lightness that corresponds to the reflectances of objects in each wave band.

It is one thing to measure a triplet of lightness equivalents using a retinex filter coupled to a photomultiplier: it is quite another for the eye to determine lightnesses in the unevenly lighted world without reference sheets of white paper. I described above the ability of an isolated receptor system — the hypersensitive system of rod vision — to classify objects correctly according to their inherent reflectivity regardless of whether the objects happened to be in a brightly or a dimly lighted region of visual space. The ability of one receptor system to work in this way makes it plausible that the other three systems of normal daytime vision possess the same ability, each system viewing the world through a broad but restricted region of the spectrum. The regions we duplicate with retinex filters. Each system forms a separate lightness image of the world. The images are not mixed but compared. The comparison of lightnesses at each area gives rise to the range of sensations we know as color.

How could the biological system generate a hierarchy and spacing of lightness values given only the flux from each point in a scene and knowing nothing about the pattern of illumination and nothing about the reflectances of objects? The scheme I am about to describe is the most general we have found that surmounts these limitations: its physiological embodiment could take many forms.

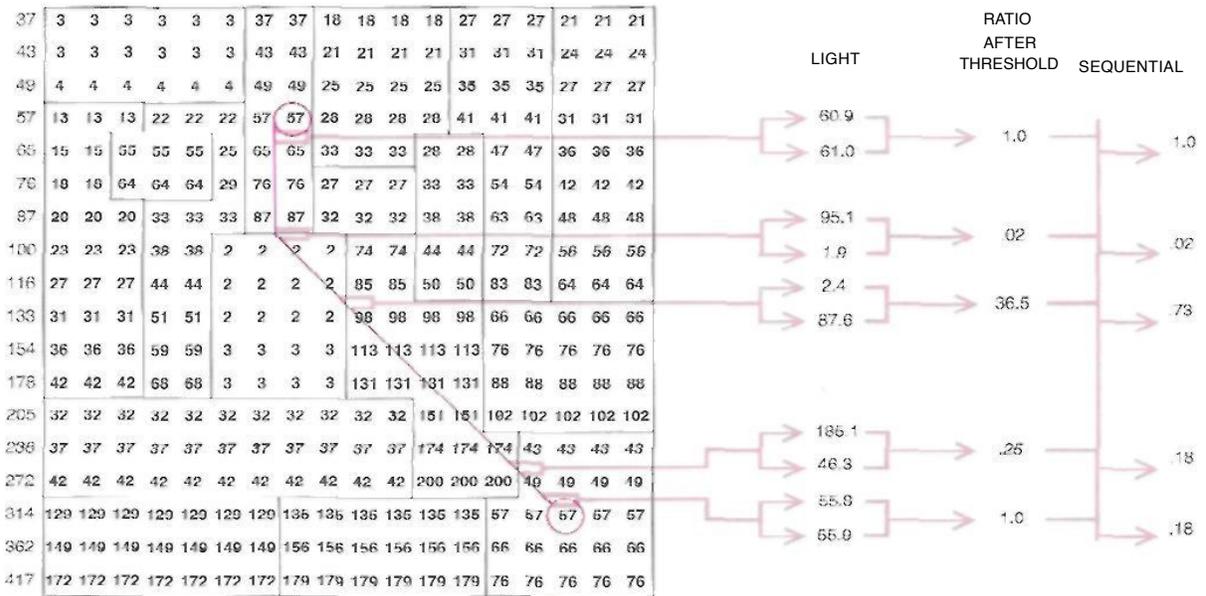
Let me begin by pointing out the significance of edges in defining objects or areas in a scene. If a sheet of white paper is lighted strongly from one side, we see no discontinuity in color from one side to the other. Let us now imagine two light detectors positioned to measure the luminance from two different places on the paper. If the illumination is nonuniform, the luminances of the two places will of course be different. As the two detectors are moved closer together the luminances approach the same value and the ratio of the two outputs approaches unity. If, however, the two detectors bridge the boundary between two areas that differ abruptly in reflectance, such as would be the case with even a pale gray square on the white paper, the ratio of the outputs of the two detectors will approach the ratio of the two reflectances. Thus the single procedure of taking the ratio between two adjacent points can both detect an edge and eliminate the effect of nonuniform illumination. If we process the entire

image in terms of the ratios of luminances at closely adjacent points, we can generate dimensionless numbers that are independent of the illumination. These numbers give the ratio of reflectances at the edge between adjacent areas: the reflectances themselves are not yet ascertained.

In order to determine reflectances we need to relate all these ratios of reflec-

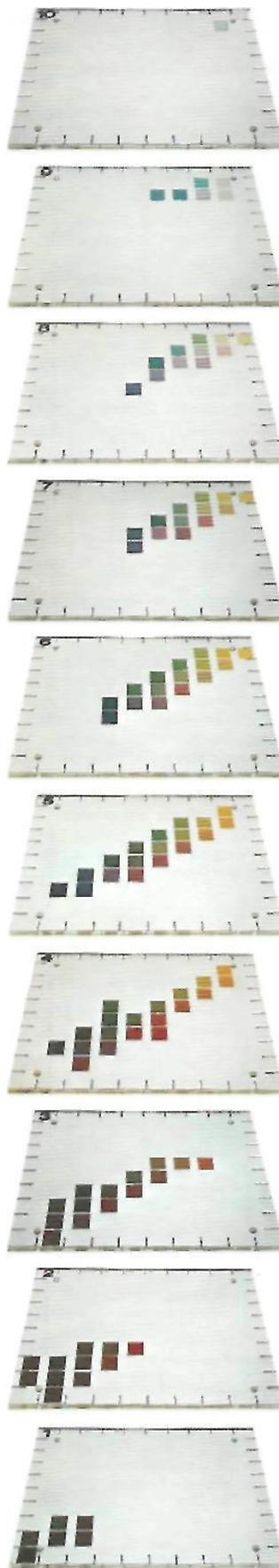
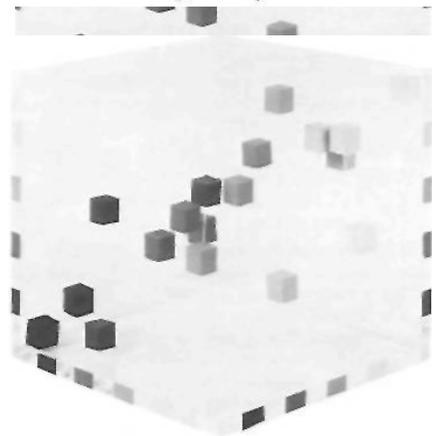
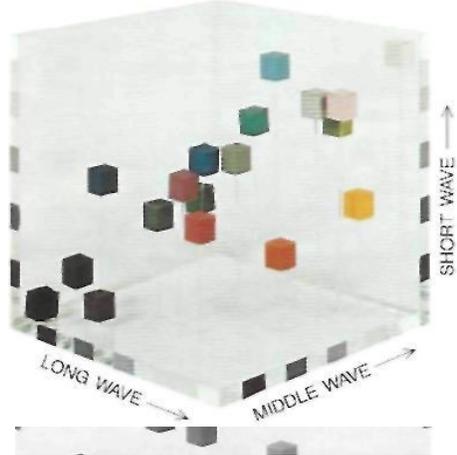
SCALED INTEGRATED REFLECTANCE OF AREAS IN MONDRIAN

AGREEMENT IN SCALED INTEGRATED REFLECTANCES between Mondrian areas and Munsell chips chosen to match them is summarized for all three wave-band systems. The scaled integrated reflectances of five Mondrian areas and matching Munsell chips were determined as is described in illustration on page 10. In this graph triplets of scaled integrated reflectances of five Mondrian areas that sent identical fluxes to the eye are plotted against scaled integrated reflectances of Munsell chips chosen to match them. Although the dots collectively represent correspondence for all three retinex wave bands, any particular dot denotes the degree of correspondence on one retinex wave band between a Mondrian area and a Munsell chip. Close correspondences show that scaled integrated reflectance is physical correlate of the sensation "lightness," showing precision with which a triplet of lightnesses determines color.



MORE REALISTIC CASE OF GRADED ILLUMINATION is handled equally well by the sequential-product method to arrive at the same reflectance value of .18 for the brown area at the end of the path, even though here the long-wave retinex system receives as much flux from the middle of the brown area (57) as it does from the middle

of the white area (57). The scheme hence provides a means for arriving at computed reflectance independent of flux and without resort to white cards as standards. Precise values of light flux along pathway in this diagram were derived from a computer program that works with 75 values between every two values printed within Mondrian.



tances in the field of view. Given the ratio of luminances at the edge between a first area and a second one, we multiply it by the ratio of luminances at the edge between the second area and a third. This product of ratios approaches the ratio of reflectances between the first and third areas, regardless of the distribution of illumination. Similarly, we can obtain the ratio of reflectances of any two areas in an image, however remote they are from each other, by multiplying the ratios of all the boundaries between the starting area and the remote area. We can also establish the ratio of the reflectance of an area on the path by tapping off the sequential product reached at that area [see illustrations on preceding page].

We are now coming close to the answer to the question: How can the eye ascertain the reflectance of an area without in effect placing a comparison standard next to the area? The sequential product can be used as a substitute for the placement of two areas adjacent to each other, thus defining a photometric operation feasible for the eye.

The remaining task is to suggest how the eye can discover the area of highest reflectance in the field of view and then decide whether that area is actually white or some other color. In the model we have proposed, sequential products are computed along many arbitrary pathways that wander through the two-dimensional array of energies on the model's "retina." Since the pathways

COLOR "SOLID" shows the location of all perceivable colors, including white and black, in a three-dimensional color space constructed according to the author's retinex theory. The position of a color in this space is determined not by the triplet of energies at a point but by the triplet of lightnesses computed by the eye for each area. The color photograph at the top left shows the location of representative colors throughout the space. The direction of increasing lightness along each axis is shown by the arrows. The three black-and-white photographs of the color solid were taken with retinex filter-film combinations. They show the lightness values of the representative colors as they would be perceived separately by the eye's long-wave (top), middle-wave (middle) and short-wave (bottom) visual pigments. The set of 10 color pictures at the right represents horizontal planes cut through the three-dimensional color space. Each plane is the locus of colors possible with a constant short-wave lightness. For example, the fifth plane from the bottom shows the variety of color sensations from all possible long- and middle-wave lightness values when those values are combined with a short-wave lightness of 5. The colored squares are samples taken from *The Munsell Book of Color*. In general the blank areas on each plane represent regions where colors could be produced only by fluorescent dyes, if they were produced at all.

can begin anywhere, not just in regions of the highest reflectance, the first value in any sequence is arbitrarily assumed to be 100 percent. Because of this deliberately adopted fiction the sequential product becomes greater than unity whenever the path reaches an area whose reflectance is higher than that of the starting area.

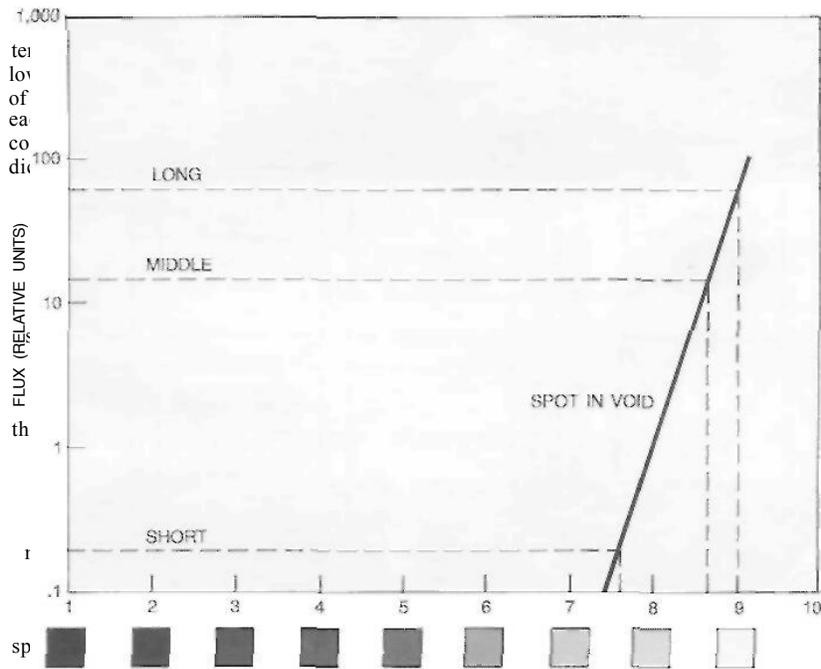
The attainment of a sequential product greater than unity indicates that the sequence should be started afresh with the new area of high reflectance taken as being 100 percent. This procedure is the heart of the technique for finding the highest reflectance in the path. After the path reaches the highest reflectance in the scene, each of the sequential products computed thereafter becomes a fraction of the highest value. A satisfactory computer program has been designed to study the number of paths, their lengths and convolutions, the threshold values for recognizing edges and, perhaps most important, how to utilize all the pathways starting in all areas.

The biological counterpart of this program is performed in undetermined parts of the pathway between the retina and the cortex. The process that corresponds to computing sequential products does not involve the averaging of areas or the averaging of flux. It does, however, call for an arithmetic that extends over the entire visual field. Furthermore, since the relevant phenomena are seen in a brief pulse of light, all the computations and conclusions about lightness must be carried out in a fraction of a second without dependence on eye movement. With a single pulse, eye movement, by definition, is not necessary. With continuous illumination the normal quick motions of the eye probably serve to maintain the freshness of the process.

With our computer model we can obtain a triplet of lightnesses for each area in the color Mondrian that corresponds closely to the lightnesses one would measure with a combined retinex filter and photomultiplier. The color corresponding to any given triplet can be visualized with the aid of the color "solid" we have built, in which the Munsell colors are located in three dimensions in "lightness-color space" according to their lightness values measured in three

wave bands through retinex filters [see illustration on page 14].

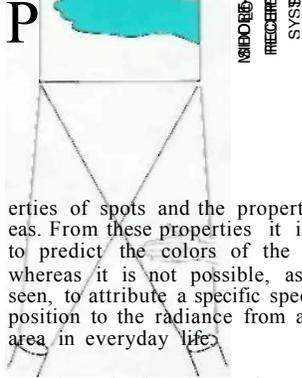
In normal images the sensation of white light will be generated by any area that is placed at the top of the lightness scale by all three retinex systems. On the other hand, an area that stands at the top of only two of the three lightness scales will be seen as some other color. Hence an area that is at the top of the lightness scale in the long- and middle-wave systems but is surpassed in lightness by some other area in the short-wave sys-



One can readily measure the flux at various gradients involving both the properties of the eye from a spot of light in a void. By

LIGHTNESS SENSATION

SPOT OF LIGHT IN A VOID, that is, a single spot of narrow-band light viewed in an otherwise totally dark environment, has a color that would seem to depend solely on its wavelength. The color can also be explained, however, by the retinex theory in terms of lightness as perceived by the eye's three receptor systems. Psychophysical measurements show that when the eye is presented with a spot of light in a void, the perceived lightness is changed only slightly by very large changes in flux, as is indicated by the straight line. For example, if the spot is composed of a narrow-wavelength band centered, say, at 600 nanometers, the three cone pigments will absorb the flux in quite different amounts because of the shape of their absorption curves. In arbitrary units the long-wave pigment might absorb 80 units, the middle-wave pigment 20 units and the short-wave pigment a few tenths of a unit at most. If these ratios are plotted on the spot-in-a-void curve, the corresponding lightness values are 9 for the long waves, 8.5 for the middle waves and 7.5 for the short. This combination of lightnesses is perceived as a light reddish orange, not ordinarily seen under normal conditions unless surfaces are fluorescent.



erties of spots and the properties of areas. From these properties it is possible to predict the colors of the spectrum, whereas it is not possible, as we have seen, to attribute a specific spectral composition to the radiance from a colored area in everyday life.

Perhaps the first observation pointedly relevant to the mechanism of color formation in images is not Newton's spectrum but the phenomenon of colored shadows, described in 1672 by Otto von Guericke. "This is how it happens," he wrote, "that in the early morning twilight a clear blue shadow can be produced upon a white piece of paper [by holding] a linger or other object... between a lighted candle and the paper beneath." This important experiment, we now know, depicts an elementary example of generating three different lightnesses on the three receptor systems. A diagram of this experiment with long-wave ("red") light and white light appears below. Here the color of the shadow is blue-green. The diagram shows that the triplet of lightnesses in the shadow corresponds to the blue-green color one would predict for it from its posi-

tion in lightness-color space.

One can now understand the red-and-white images of our early work as a procedure that carries the colored shadow to a richly variegated family of colors no longer in shadows but in images. The colors seen in a red-and-white projection can be readily predicted by extending the analysis followed in predicting the color of von Guericke's shadow. To demonstrate this point we reproduce on page 17 the "red" and "green" separation images used in making a red-and-white multicolored projection. (In this demonstration no attempt is made to reproduce the colors seen in the actual multicolored image.) The red-and-white projection was photographed through long-, middle- and short-wave retinex film-filter combinations. The three images are reproduced below the pair of long- and middle-wave separation images that were superposed to make the red-and-white image. The significant point is that when the eye views the red-and-white images on the screen with its own retinex system, it is provided with a triplet of lightnesses for each part of the scene that resembles the triplet it would obtain if it viewed the original scene di-

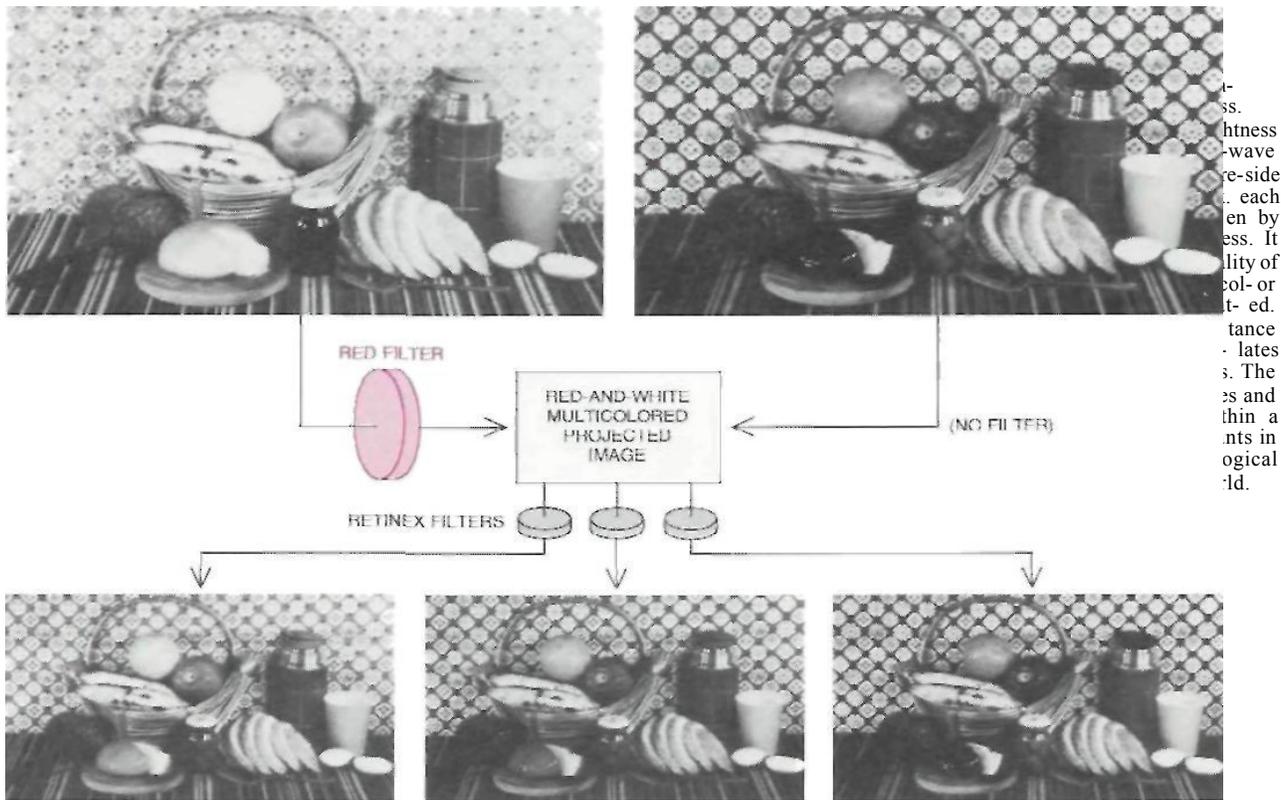
rectly. In this important meeting point of the blue-green shadows with the colored images, provided by the red-and-white display, the extended taking and multiplication of ratios determine the lightness of each small area. Finally, all these principles are applied in everyday ternary vision, which creates a distinct lightness image for each of the three sensitive systems and compares them in order to generate color.

The train of interlocking concepts and experiments started 25 years ago with the observation that the relative energies of the red-and-white projectors can be altered without changing the names of the various colors. This observation negated the simplistic explanation in terms of contrast, fatigue and surround and led to the fundamental concept of independent long- and short-wave image forming systems that ultimately evolved to the concept of three independent retinex systems and to the Mondrian demonstration. The concept of the percentage of available light on each wave band as a determining variable and the technique of measuring it evolved to the concept that lightnesses

		ADSORBED FLUX (ARBITRARY UNITS)				
SHADOW		FROM "WHITE"	FROM TWO	LIGHTNESS	COLOR	
		ILLUMINATOR	ILLUMINATORS			
		(SCATTER)	COMBINED			
		100				
		100				
"WHITE"		100				
ILLUMINATOR	LONGWAVE					
	("RED")					
	ILLUMINATOR					

BLUE-GREEN COLORED SHADOW is seen when a hand or some other object is placed in the beam of a projector that is sending long-wave ("red") light to a screen while the screen is illuminated by a beam of white light. The author regards Otto von Guericke's description in 1672 of seeing colored shadows made by candlelight as the first observation pointedly relevant to the mechanism of image and color formation. In the analysis at the right it is assumed that one projector sends white light to the screen. The other projector, equipped with a red filter, sends only long wavelengths to the screen. Assume that the white light contributes 100 arbitrary units of flux to each of the short-, middle- and long-wave receptors. The long-wave flux is absorbed by the three receptor systems in different proportions: 100 units are

absorbed by the long-wave system, SO by the middle-wave system and five by the short-wave system. (A small amount of scattered long-wave light also appears in the shadow.) The third column of boxes shows the combined amounts of flux from both sources absorbed by each receptor system. The fractions represent the ratio at edges of the flux from within the shadow divided by the flux from outside. The fourth column shows the lightness on each receptor system. The lightness of the lightest place in the scene for each receptor system will be near the top of the lightness scale, being determined by the flux of radiant energy in the same way that a spot has its lightness determined by flux. Triplet of lightnesses within the shadow falls in the region of color space that the eye perceives as being blue-green.



t-
ss.
htness
-wave
re-side
... each
en by
ess. It
lity of
col- or
t- ed.
tance
- lates
s. The
es and
thin a
nts in
ogical
ld.

RETINEX RECORDS OF RED- AND- WHITE projections show that red-and-white images produce a triplet of lightnesses for each part of the scene that are consistent with the observed color sensations. The two photographs in the top half of this illustration are reproductions of the long-wave (*left*) and middle-wave (*right*) separation records taken of the original still life. The long-wave record was projected onto a screen with a long-wave (red) filter in the beam of light. The middle-wave record was projected in superposition onto the same screen in the light of a tungsten-filament lamp. Three retinex photographs were then taken of projected images appearing on

screen. The retinex records are reproduced in the bottom part of the illustration: long-wave at the left, middle-wave in the middle and short-wave at the right. The colors seen in red-and-white projections are those one would expect from their triplets of lightnesses. The apple is light on the long record and darker in the middle and short records. The orange is lightest on the long record, intermediate on the middle record and darkest on the short. It is impressive that with his own retinex systems the observer can see a blue cup, a brown straw basket and pale yellow bananas with lightness differences so small as to challenge photoengraving process used to reproduce photographs.

The Author

EDWIN H. LAND is chairman of the board, director of research and chief executive officer of the Polaroid Corporation. Born in 1909, he attended Harvard College, where he developed a new type of polarizing filter in the form of an extensive synthetic sheet. In 1937 he founded Polaroid for research in the new field of applied polarization, and in 1944 he began his pioneering work in the development of "instant" photography. His one-step photographic process was first demonstrated to the Optical Society of America in February, 1947, and was made available to the public at the end of 1948. Land has received 14 honorary degrees, has held visiting academic appointments at Harvard and is currently Institute Professor (Visiting) at the Massachusetts Institute of Technology. From 1960 to 1973 he was consultant-at-large to the President's Science Advisory Committee, and in 1967 he received the National Medal of Science. This year, on the occasion of his 500th U. S. patent, he was elected to the National Inventors Hall of Fame. Land has pursued his lively interest in the mechanisms of color vision for the past 25 years.

Bibliography

- COLOR VISION AND THE NATURAL IMAGE: PART I. Edwin H. Land in *Proceedings of the National Academy of Sciences*, Vol. 45, No. 1. pages 115-129: January, 1959.
- COLOR VISION AND THE NATURAL IMAGE: PART II. Edwin H. Land in *Proceedings of the National Academy of Sciences*, Vol. 45, No. 4. pages 636 - 644: April, 1959.

- INTERACTION OF THE LONG-WAVE CONES AND THE RODS TO PRODUCE COLOR SENSATIONS. John J. McCann and Jeanne L. Benton in *Journal of the Optical Society of America*, Vol. 59, No. 1. pages 103-107: January, 1969.
- LIGHTNESS AND RETINEX THEORY. Edwin H. Land and John J. McCann in *Journal of the Optical Society of America*, Vol. 61. No. 1. pages 1-11; January, 1971.

The Cover

The pattern on the cover was used in experiments testing Edwin H. Land's retinex theory of color vision. Because the pattern bears a resemblance to the works of the Dutch painter Piet Mondrian. Land refers to this display and similar ones as Mondrians. In more elaborate examples (*see top illustration on page 5*) perhaps 100 pieces of paper of various colors and sizes are mounted on large boards and so arranged that each piece of paper is surrounded by at least five or six other pieces of different colors. In a typical demonstration the Mondrian is illuminated by projectors that provide adjustable amounts of radiant energy in three wave bands: long ("red"), middle ("green") and short ("blue"). With the proper selection of the mixture of illuminants falling on the Mondrian the radiant flux reaching the eye from any selected area can be made to match the flux that had previously reached the eye from a totally different area. In the first instance the selected area could have been red: in the second instance it could have been green. With the same flux of energy reaching the eye the two areas will still be seen as red and green. (Cover photograph by Julius J. Scarpetti)