5 Seeing Colour

John Mollon

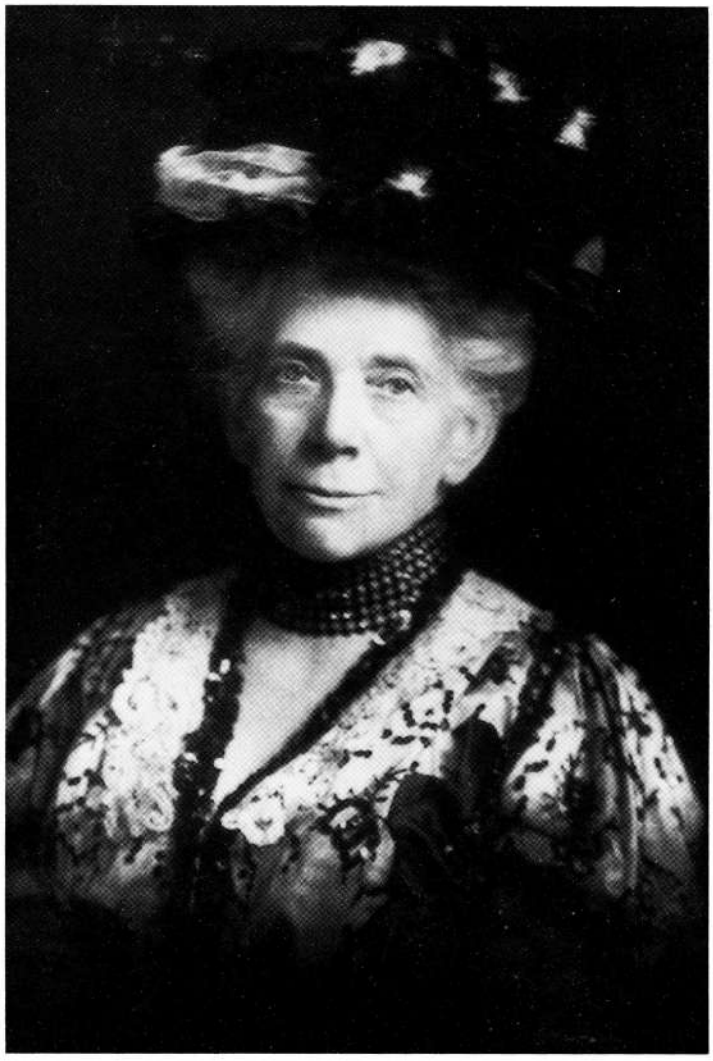

Figure 1 Christine Ladd-Franklin (1847–1930), pioneer American feminist, who put forward an evolutionary account of human colour vision. Ferdinand Hamburger, Jr, Archives of the Johns Hopkins University.

Even the Preface to her collected papers describes her as belligerent. She graduated from Vassar while Darwin was still in his prime, and to the

time of her death in 1930 she robustly defended the evolutionary theory of colour perception that she first put forward in 1892. Without the advantage of molecular biology, Mrs Christine Ladd-Franklin (Figure 1) surmised that the light-absorbing molecule of the retina had undergone successive differentiations, to give first achromatic, then dichromatic, and finally trichromatic vision. I believe, with Mrs Ladd-Franklin, that to understand our own colour vision we must understand how it came to be the way it is. I shall argue that it depends on two distinct subsystems, a relatively recent one overlaid on a phylogenetically ancient one. Denis Baylor (Chapter 4) has already introduced the idea that our rich experience of hue depends on the proportions of light absorbed in just three classes of cone cell in our retinae (see Figure 2(*c*)). What I should like to do here is to emphasise that the three types of cone are not equal members of a trichromatic scheme, but rather that they evolved at different times for different purposes.

The ancient subsystem of colour vision

For most diurnal mammals the main business of vision (the discrimination of position, form and movement) depends on a single class of cone with its peak sensitivity in the middle of the spectrum. Recall, however, from Chapter 4 that any individual class of cone is colour blind and so we need to be able to compare the photon absorptions in at least two classes of cone before we can discriminate colour. A long time ago – probably before the mammals evolved – a second population of cones, sensitive to short-wave (violet) light, were added to the array of light-sensitive cells: a comparison of their signals with those of the predominant class of cones provided the ancient system of colour vision (Figure 2(*a*)). In the typical mammalian retina, these violet-sensitive cones are sparsely scattered within the light-sensitive array, and in primates and human beings they represent only a tiny minority of all cones. Using the technique of microspectrophotometry, in which a 2 μm wide beam of light is passed through individual cone cells, James Bowmaker and I have recently been able to measure the absorption spectrum of each cone in small patches of intact retina. In the central region of the retina, the fovea, we find that only about 3% of all cones are of the short-wave kind,

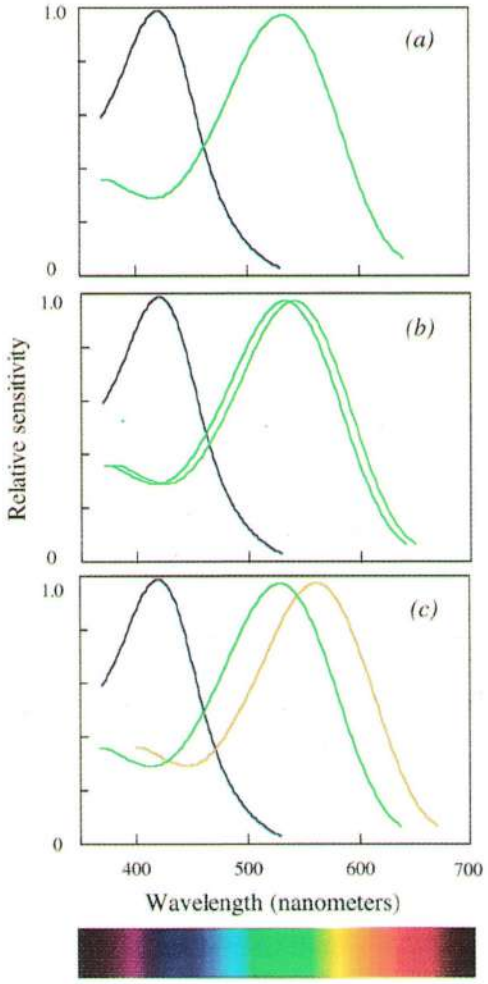

Figure 2 The light absorption curves of pigments in the cone cells of the eye. Each panel shows the relative proportion of light absorbed by a given photopigment at different wavelengths.

(*a*) Dichromacy. This form of colour vision is found in many non-primate mammals, in male New World monkeys, and in about 2% of human males. It depends on comparing the relative light absorption in violet-sensitive cones with the absorption in more numerous cones that have peak sensitivity in the green or green–yellow part of the spectrum.

(*b*) Anomalous trichromacy. This form of colour vision is found in about 6% of human males: a limited discrimination is achieved in the red–green range by comparing the absorptions in two pigments that have very similar positions in the spectrum. Some female New World monkeys exhibit similar sets of photopigments.

(*c*) Normal trichromacy. This panel shows the arrangement of three photopigments found in Old World monkeys, in apes, and in those human subjects with normal vision. The exact spectral positions of the pigments vary slightly between species and between individuals.

and in a small area at the very centre they may be completely absent.

So the ancient mammalian system of colour discrimination depends on a comparison of the rates at which photons are absorbed in the sparse short-wave cones and in the more numerous type of cone that has its peak sensitivity in the middle of the spectrum. The comparison is carried out by nerve cells within the retina that draw inputs of opposite sign – excitatory or inhibitory – from the two classes of cone cell (Figure 3(*a*)). It is this subsystem of colour vision that is commonly preserved in those of our own species whom we label 'colour-blind'. Phenomenologically, in the case of both the colour-blind and the colour normal, the ancient subsystem divides colours into 'warm' and 'cold': stimuli appear warm if long wavelengths predominate and cold if short wavelengths predominate. A mother will find that she can communicate unambiguously with her colour-blind son if she speaks only in terms of this warm–cold dimension.

Why are the short-wave cones so sparse? The answer is probably chromatic aberration, one of the several optical imperfections of the eye. Von Helmholtz once said that, if an instrument maker sent him the human eye, he would send it back. (There are actually rather few customer complaints, and we may guess that this is because the retinal image is the one optical image never intended to be looked at; for the brain's purpose is to reconstruct the *external* world, not to show us the distorted and degraded image on the retina.) Chromatic aberration arises from the fact that blue rays are more refrangible than red. The two ends of the spectrum cannot be concurrently focused by the human eye. The eye normally focuses for the yellow light that is so abundant in our world and so the short-wave component of the retinal image is permanently out of focus. The visual system would gain little by sampling the blue and violet component of the image at higher density. The Great Instrument Maker provides us with short-wave cones primarily for colour vision. We make little use of them for many of the functions of the eye, such as the discrimination of fine detail or the detection of movement and flicker.

It is easy to demonstrate the poor spatial resolution of the ancient subsystem of colour vision. Take a blank transparency for an overhead projector and, with a suitable felt-tipped pen, tint one half of it a pale lime yellow (Figure 4). The boundary of the yellow-coloured area should be sharp and straight. When the transparency is projected with an overhead

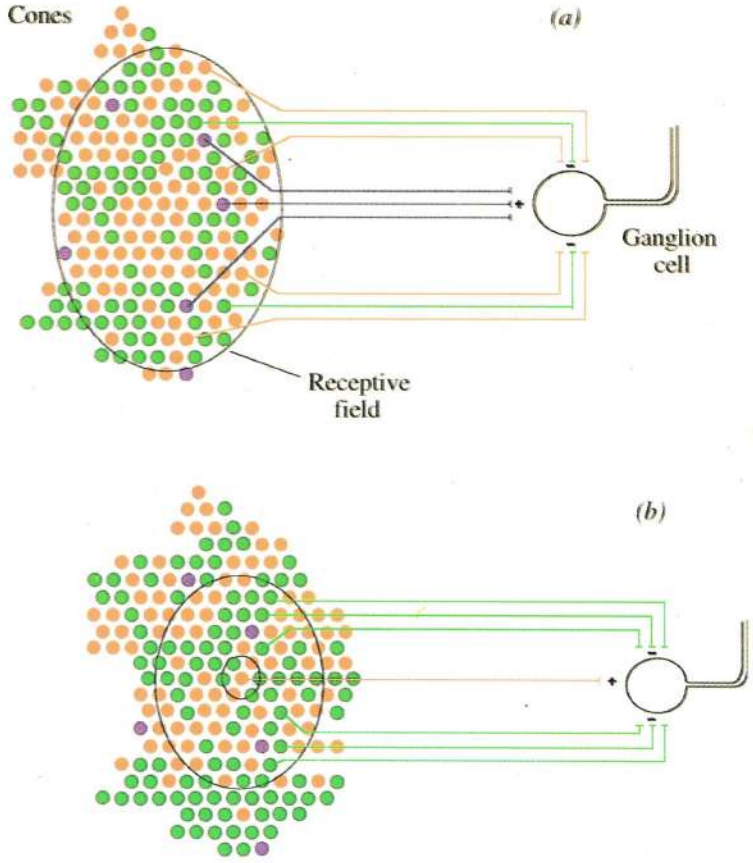

Figure 3 Ancient and modern subsystems of colour vision. (*a*) The phylogenetically older subsystem. A retinal ganglion cell draws excitatory (+) inputs from the short-wave cones and inhibitory (–) inputs from the long- and middle-wave cones in a small local region of the photoreceptor array. When this small region (the receptive field of the ganglion cell) is illuminated with homogeneous white light, the ganglion cell will not respond, since its excitatory and inhibitory inputs are in balance; but it will respond vigorously when blue or violet light falls on the receptive field. Short-, middle- and long-wave cones are represented by violet, green and yellow discs, respectively, and the spatial distribution of the different types follows that of Mollon and Bowmaker.

(*b*) The phylogenetically recent subsystem. A ganglion cell draws opposed inputs from the long- and middle-wave cones. The excitatory and inhibitory areas of the receptive field are spatially concentric, and in the foveal region of the retina the centre input to the ganglion cell may be drawn from only a single cone.

The reader should envisage that the photoreceptor array is analysed in parallel by many ganglion cells of the two types, each cell having its local receptive field. References to 'subsystem' in the main text refer to all the cells of a given type.

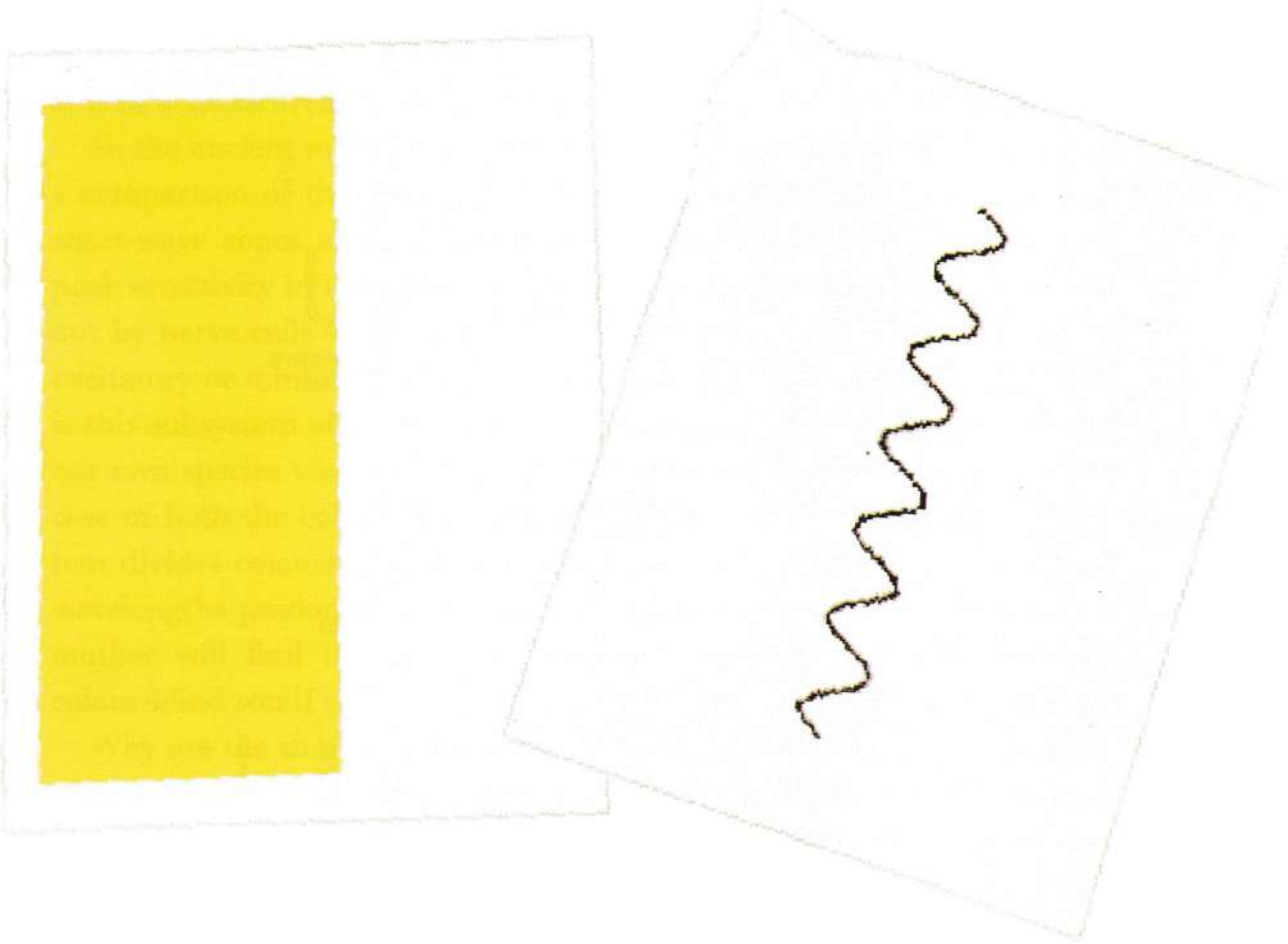

Figure 4 Two transparencies for demonstrating the poor spatial resolution of the older subsystem of colour vision. When the two are superposed using an overhead projector, with the sinusoidal line coinciding with the central edge of the yellow–green region, the observer sees the colour follow the concavities and convexities of the superposed black line. The yellow–green should be as light as possible so as to minimise the brightness difference at its edge. This demonstration was designed by R. M. Boynton.

projector, the edge between the two areas will be visible mainly to the short-wave, violet-sensitive cones: these cones respond more strongly to the white region than to the yellow, whereas the other cones signal little difference between the two regions. Now prepare a second transparency with a sinusoidal black line running down the middle. When this wavy line is superposed on the yellow-white edge, the colours will appear to follow the black line, the yellow spreading into the positive-going excursions to the one side and the white passing into the positive-going excursions to the other. The colour signals from the short-wave cones give only

inexact information about spatial position, and it seems that signals from the more abundant longer-wavelength cones are used to decide the exact position of an edge.

The modern subsystem of colour vision

About thirty million years ago, a second subsystem of colour vision evolved in our primate ancestors. This development, like so much of evolution, arose through the duplication and modification of an existing gene. The ancestral gene was located on the X-chromosome and it coded for the light-sensitive pigment found in the main class of mammalian cone, having peak sensitivity in the middle of the spectrum. It now gave rise to two genes, with slightly different DNA sequences, which coded for two photopigments with peak sensitivities in the green and yellow-green regions of the spectrum (Figure 2(*c*)). The pigment with peak sensitivity in the green is conventionally known as the 'middle-wave' pigment; that with a peak in the yellow-green is known as the 'long-wave' pigment. The two pigments are segregated in different cone cells, and our second subsystem of colour vision depends on a neural comparison of the light absorption in these two kinds of cone. This newer subsystem allows us to make discriminations in the red–yellow–green region of the spectrum, and, because both types of cone are present in high numbers, we enjoy better spatial resolution for this dimension of colour experience. Notice that neither of the pigments has a peak in the red: a red light is simply one that produces a high ratio of long-wave to middle-wave cone signals.

Diurnal birds and many surface-dwelling fish exhibit good colour vision. But amongst the mammals, we share our own full trichromatic vision only with the Old World monkeys and the apes. It increasingly seems that the second subsystem co-evolved with a class of tropical trees characterised by fruits that weigh between 5 and 50 g, are too large to be taken by birds, and are yellow or orange in colour when ripe. The tree offers a signal that is salient only to agents with trichromatic vision. This hypothesis of co-evolution of the tree's signal and the monkey's vision can be traced back to nineteenth-century naturalists, but it takes on new plausibility in the light of recent ecological evidence that there are many species of tropical trees that are dispersed exclusively by monkeys, and

these are the trees with yellow or orange fruit (for example, several members of the family Sapotaceae). Conversely, some species of arboreal monkey, such as guenons, rely on fruit for up to 85% of their diet. The tree offers a colour signal that is visible to the monkey against the masking foliage of the forest, and in return the monkey either spits out the undamaged seed at a distance or defecates it together with fertiliser.

In short, monkeys are to coloured fruit what bees are to flowers. With only a little exaggeration, one could say that our trichromatic colour vision – if not the entire primate lineage – is a device invented by certain fruiting trees in order to propagate themselves. Certainly, if we are to understand the regeneration of rain forest, one of the many biological contracts we must understand is that between fruit signals and primate colour vision.

The polymorphic colour vision of New World monkeys

One unsolved puzzle concerns the New World monkeys, which diverged from our own ancestors some 30 million years ago and which have found their own, very remarkable, route to trichromacy. Within a single species of New World monkey, such as the squirrel monkey or the marmoset, colour vision is polymorphic, that is to say there exist in the population several genetically different forms, apparently in stable equilibrium. There may be three types of male and six types of female within a species. The males are always dichromats, resembling colour-blind men in having only two classes of cone (Figure 2(*a*)) and differing only in the exact spectral position of the pigment with the longer wavelength, whereas a majority of females are trichromats, able to discriminate well in the red–green region of the spectrum. The genetic basis of this variation is instructive and may have analogues in other physiological systems. Like most mammals, the New World monkeys have only a single X-chromosome locus for a pigment in the midde-/long-wave region, but three or more versions of the gene are present in equilibrium in the population, the different versions differing slightly in their DNA sequence and coding for pigments that peak at different positions in the spectrum. A female monkey, having two X-chromosomes, has a good chance of being heterozygous, that is inheriting different versions of the gene from

her two parents; since only one X-chromosome is expressed in any given cell of a female mammal's body, she will thus come to have three types of cone in her retina – the two types of cone with peak sensitivity in the green–yellow range plus the violet-sensitive cones that she shares with her male conspecifics. Apparently the visual system of such heterozygous females is plastic enough to take advantage of the signals from the extra cones so as to gain behavioural trichromacy.

This is a particularly pure case of what population geneticists call 'heterozygous advantage': it matters only that the female inherits different versions of the gene from her parents, not which particular ones she inherits. (A recent report suggests that heterozygosity for a related molecule – one of the receptors for dopamine – is associated with reduced susceptibility to schizophrenia.) And there is a moral for those who study the inheritance of traits such as intelligence: here we have a trait, the ability to discriminate red from green, where the variance between individuals is almost all of genetic origin and yet we should expect no correlation between the abilities of parents and offspring. Needless to say, such a remarkable paradox does need some technical qualification. For the statement to be true, it is necessary first that the different versions of the gene are equally frequent in the population; and, secondly, that mating must be random with respect to this gene (for example, females must not be able to distinguish, and favour, males carrying a different version of the gene.)

Traditional field studies of New World monkeys had not given any hint of their polymorphism of colour vision, which emerged from laboratory studies in the 1980s. But the laboratory studies must now feed back to research in the rainforest, for only ecologists in the field are likely to discover why the striking variations between individuals and between sexes are maintained. The biological contract between trees and their primate dispersers is different in the Old World and the New. We do not understand why.

Individual differences in human colour vision

The gene duplication that occurred within the Old World primate lineage remains a source of recombinant mischief for a significant fraction of our

own population – those 8% of men who are colour deficient. Almost invariably, it is the newer subsystem that is altered in colour deficiency. The two very similar genes on the X-chromosome have remained juxtaposed, and so are liable to misalign when corresponding chromosomes are brought together at the stage of meiosis during the formation of the ovum. This is the stage when crossing-over occurs; that is, when DNA is exchanged between paired chromosomes. If the chromosomes are locally misaligned and crossing-over occurs in this region, hybrid genes may be formed or a gene may be lost altogether from one of the chromosomes. About 2% of men are dichromats, lacking either the long-wave pigment or the middle-wave pigment. More common, affecting 6% of men, is the milder condition called anomalous trichromacy, in which three pigments are present in the retina, but one of them is displaced in its position in the spectrum (Figure 2(*b*)).

It is instructive that picking fruit is one of the few tasks in which colour-deficient men are severely handicapped. In one of the earliest reports of colour blindness, in the *Philosophical Transactions of the Royal Society* for 1777, Huddart wrote of the shoemaker Harris, '*Large objects he could see as well as other persons; and even the smaller ones if they were not enveloped in other things, as in the case of cherries among the leaves.*' We especially need colour vision when the background is dappled and variegated; that is, when we cannot use form or lightness to find our target. For urban humans, colour vision is a luxury; but for frugivorous primates, it may be a necessity.

So far, I have referred to normal colour vision as if it were a single condition. But of especial interest is the recent realisation that the modern subsystem varies amongst those of us whose colour vision is nominally normal. You and I may live out our lives in slightly different perceptual worlds. Although we may both enjoy colour vision that is officially normal, coloured objects that look alike to you may look distinctly different to me, and those that look different to you may look identical to me. This small but irreducible discrepancy in our sensations can now be traced to the minimal possible genetic difference; that is, to a single nucleotide difference in our X-chromosomes. Let me elaborate.

Denis Baylor (Chapter 4) has described the structure of the protein molecules on which all our vision depends. They are members of the

much larger superfamily of heptahelical receptors, which includes the molecules found in the sensory cells of the nose as well as many of the molecules that recognise the chemical signals between nerve cells. Figure 5 represents photopigment molecules embedded in the multiply enfolded membrane of a retinal cone cell. Each molecule consists of seven helices that span the membrane of the cell. By collating the genes and the photopigments of individual New World monkeys, it has been possible to

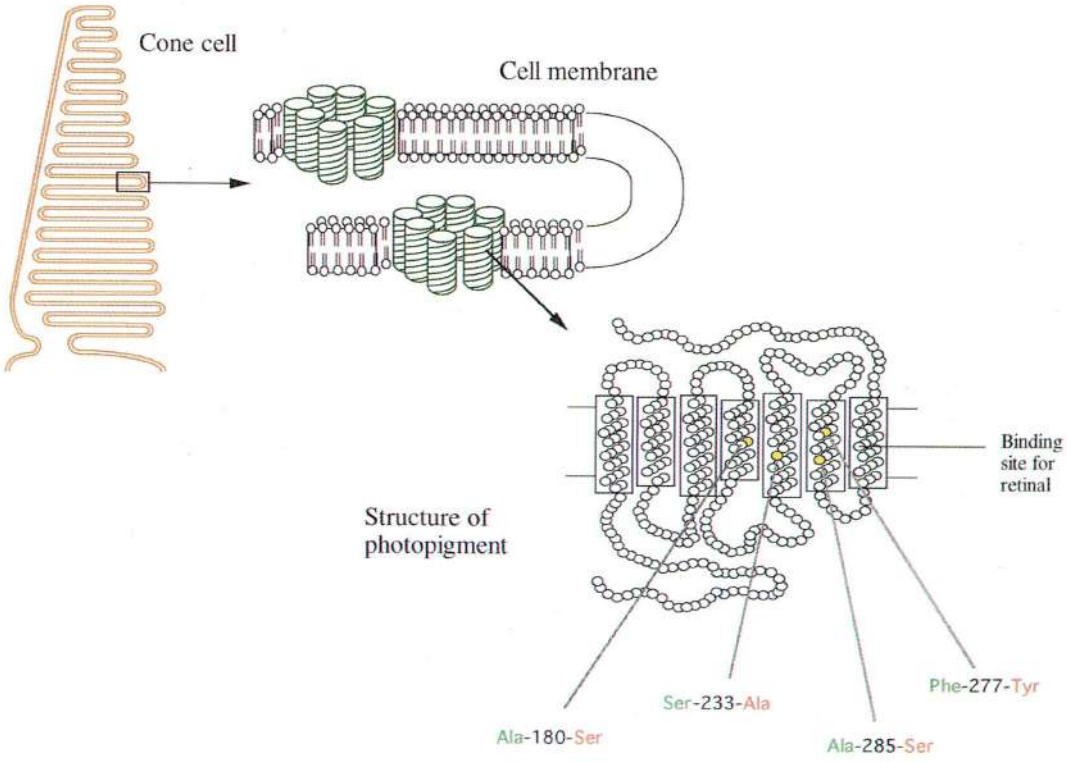

Figure 5 The enfolded membrane of a cone cell (top left) is packed with photopigment molecules. Each of these molecules consists of seven helices, which span the membrane of the cell and are linked by loops outside the membrane. Represented at the bottom right is the sequence of amino acids that makes up this heptahelical molecule. Indicated in the seventh helix is the amino acid (lysine) that binds to the chromophore (retinal) that lends the molecule its spectral absorption. Highlighted in yellow, in the fourth, fifth and sixth helices, are four amino acid residues (numbers 180, 233, 277 and 285 in the sequence) that are known to influence the exact wavelength to which the pigment is maximally sensitive. The codes below represent the alternative amino acid residues at these four positions: in each case the amino acid indicated in green is the alternative that shifts the peak sensitivity of the molecule to shorter wavelengths, that in red the one that shifts sensitivity to longer wavelengths.

identify the surprisingly small number of amino acid residues that make the difference between the long-and middle-wave pigments. Of particular interest is the 180th amino acid residue in the chain, for this is the one that commonly varies in the long-wave pigment in the normal human population. If, on a man's X-chromosome, the long-wave gene codes for the amino acid serine at position 180, then the peak sensitivity of his long-wave pigment will lie at a longer wavelength than if he inherits the code for alanine at this position. Groups of scientists in Seattle and Milwaulkee have shown that this difference is reflected in the settings that different individuals make in a 'Rayleigh match': when men are asked to adjust a mixture of red and green light to match a pure, monochromatic, orange light, those with serine at position 180 will tend to choose a mixture with less red in it than will those with alanine. A white male population divides in proportions of about 60 : 40.

In years to come I think we shall look back on this finding as the first example where a normal variation in our mental worlds is traceable to a polymorphism – a normal variation – in our genes. In this case we know almost all the steps in the chain: we know how the difference of a single nucleotide in the DNA changes the protein that is coded for, we know how that alters the spectral sensitivity of the resulting photopigment, and we know how it will alter neural signals in the visual pathway. Within the next ten years we shall surely know of many cases where differences in our perceptual, cognitive and emotional worlds are traceable to analogous genetic variations.

At least with respect to the present issue, women are more complicated than men, since they have two X-chromosomes and thus two sets of genes for the long-and middle-wave photopigments.

What happens if – as is statistically very likely – a woman inherits from her parents two different versions of the long-wave gene or two different versions of the middle-wave gene? Do the New World monkeys offer us a model? In the case of squirrel monkeys, a basically dichromatic species, the heterozygous females become trichromatic. Is it possible that in our basically *tri*chromatic species, a subset of females become tetrachromatic, enjoying an extra dimension of colour experience? The best candidates might be women who are carriers of the mild forms of anomalous trichromacy and whose affected sons must achieve their residual

red–green discrimination by means of an abnormal pair of middle-/long-wave pigments. Hitherto, it has been assumed that such carriers simply share a little in the disability of their sons, exhibiting very mild colour anomalies, rather as the mothers of haemophiliacs have blood that clots more slowly than that of other women. But the mother of an anomalous trichromat must carry at least four genes for cone pigments – the short-wave pigment, and three different pigments in the middle-/long-wave range. She may for example carry on one of her X-chromosomes the gene for one long-wave pigment and on the other the gene for a different long-wave pigment, one shifted in its position in the spectrum. Owing, then, to X-chromosome inactivation, the retina of such a woman should contain four classes of cone, because the two alternative forms of the middle-wave pigment will be manufactured in different classes of cone. Can she take advantage of them to gain tetrachromatic vision?

Gabriele Jordan and I have been recently made an experimental search for tetrachromatic women. Our subjects included women who were proven carriers (in that they had sons with various forms of colour deficiency and anomaly) as well as women who had normal sons. We used a computer-controlled colour-mixing apparatus to study colour matching in a part of the spectrum where the normal, trichromatic, observer is effectively dichromatic – can make colour matches with only two variables. This is the spectral region from green to red, 550 nm to 690 nm, where the short-wave cones are quite insensitive and where a normal trichromat can match any wavelength with some unique mixture of red and green. What we did was to add an additional variable. We offered our subjects two mixtures. The first mixture (A) was of a red and a yellow and the second (B) was of a green and a deep orange. Each of these mixtures is capable of giving a range of oranges, from a reddish orange to a yellowish orange (Figure 6). The two mixtures were alternated every two seconds and the subject was asked to adjust them until they matched. In successive trials, the computer randomized the starting points for the two mixtures, and the subjects were instructed to find as large a range of matches as they could.

The normal woman, being dichromatic in this spectral region, should be able to make a whole range of matches in the colour space of our

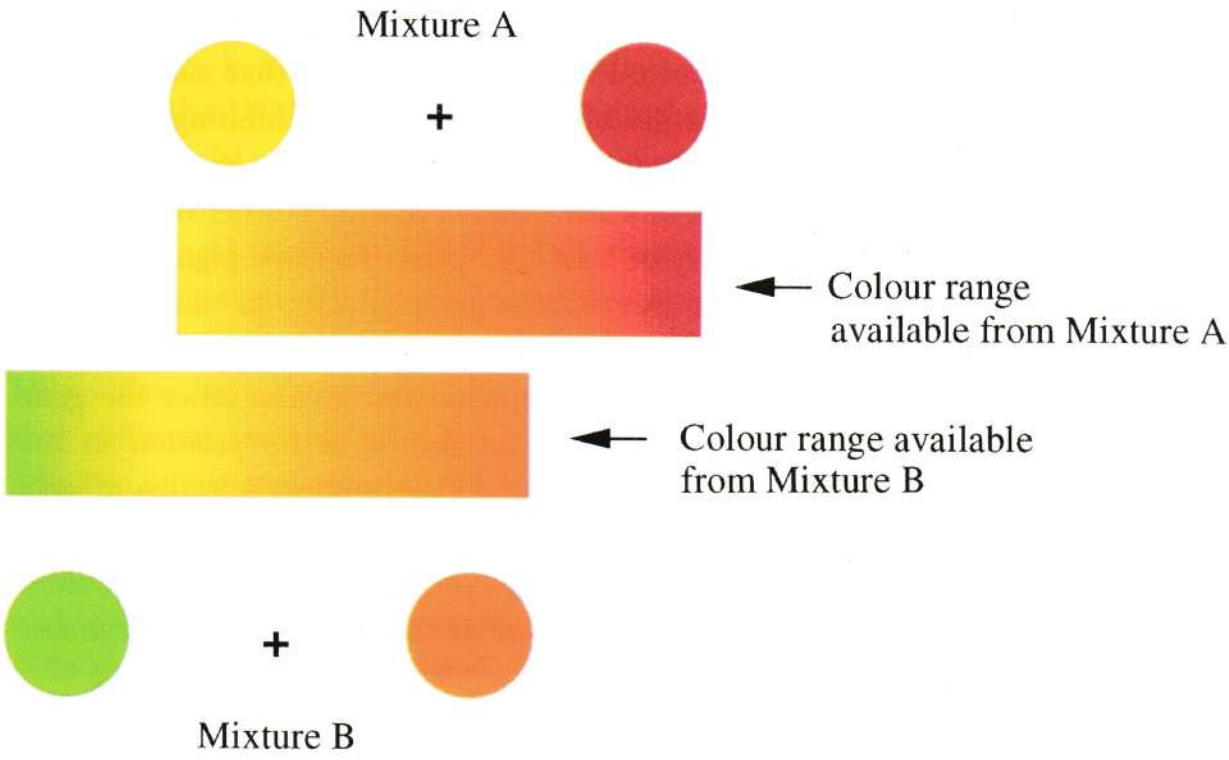

Figure 6 A test for tetrachromacy. For the normal observer, Mixture A gives a range of colours from yellow to red, while Mixture B gives a range from green to reddish orange, and so there should be a range of pairs of mixtures that match – mixtures that give colours lying between yellow and deep orange. But for a tetrachromat most of these physically different pairs should look different and only one unique match should be possible.

experiment: there should for her be a range of colours, running from a yellow orange to a reddish orange, that look the same whether they are made by mixing red and yellow or by mixing green and deep orange. But a tetrachromat, having an extra independent signal in this spectral region, would be expected to find only one mixture of the red and yellow that matched a mixture of green and deep orange. In fact, if tetrachromatic women exist, then they are rarer than would be expected from the analogy with New World monkeys. Although our experiments are still in progress and occasional carriers show a very limited range of matches, we are satisfied that the majority of carriers of anomalous trichromacy are not tetrachromats by this test and so the interesting question arises of what it is that allows the female monkey to take advantage of her extra class of cone when most heterozygous women cannot.

Perceptual organisation and the two subsystems

I have argued above that one advantage of colour vision is that it allows us to detect a spectrally distinct target (e.g. a ripe fruit) against a background that varies randomly in lightness and form. A second role for both the ancient and the modern colour systems lies in perceptual organization: colour helps us to identify which elements in a complex visual field belong together, i.e. make up a common object. This capacity of colour to link disparate components of the scene has been well recognised and exploited by painters (see Bomford, Chapter 1, and Riley, Chapter 2), and is today used and misused by those who design screens of information in teletext and other systems.

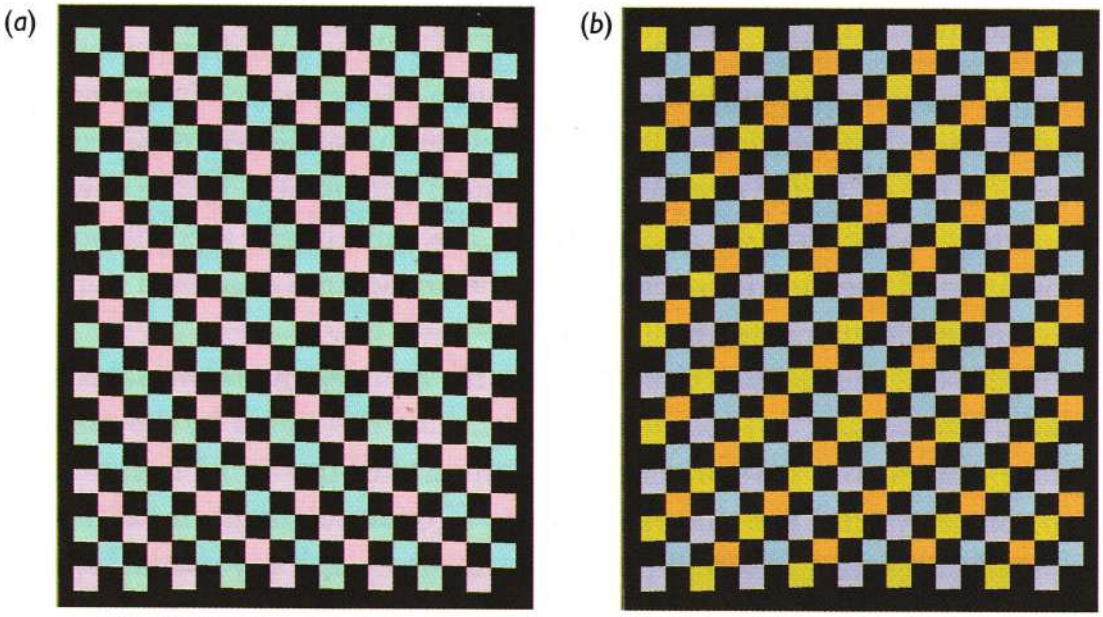

Figure 7 The role of colour in perceptual organisation. In both panels the older colour signal is modulated along the diagonal from lower right to upper left and the newer colour signal is modulated from lower left to upper right. In (*a*), where the modulation of the new signal is greater, the reader with normal colour vision will see the array predominantly running from lower right to upper left, while in (*b*), where the modulation of the older colour signal is predominant, the opposite organisation presents itself. Owing to the limitations of colour reproduction, the chromaticities and luminances shown here do not correspond exactly to those used in the experiments described in the text.

But what is the relative organising power of the two subsystems? Benedict Regan and I have recently been tackling this question experimentally. Our displays look like the arrays in Figure 7. The subject's task is to report whether the predominant organisation of the array runs from lower left to upper right or from lower right to upper left; and we arrange that the old and the new subsystems impose contrary organisations. In Figure 7(*a*) it is the new subsystem that is dominant, in Figure 7(*b*) the old. These are abstract arrays, but if we translate the discourse into that of the painter, it will be clear that this is the same organising power of colour as Bridget Riley discusses in Chapter 2.

We vary the signal strength of a given subsystem by varying the colour difference between the alternating lines of the array. In our formal experiments, the computer randomised from presentation to presentation which colour subsystem was associated with which orientation. For a series of colour differences on one of the two chromatic dimensions, the computer adjusted the depth of modulation of the other dimensions to find the point at which on 50% of occasions the subjective organisation was determined by the ancient colour system and on 50% by the new. An important factor that we also varied in the experiment was the spatial separation of the individual patches in the array, from about ten minutes of visual angle to one degree.

Figure 8(*a*) shows some typical results for a subject with normal colour vision. The vertical axis represents the size of the signal in the ancient colour system, and the horizontal axis, the size of the signal in the modern colour system. The different subsets of data points are for different separations of the individual patches in the array. For any spatial separation, the lines represent the balance points of where the two colour signals have equal organising power; they show how much modulation of the ancient system is equivalent to a given modulation of the modern system. For any one spatial separation, the potency of the two signals grows concomitantly, the data points fall on straight lines. But spatial separation has a dramatic effect: when the patches are close together, a large modulation of the ancient system is required to balance a small modulation of the new system, but when the patches are further apart the ancient colour system, the warm–cold dimension, becomes more potent.

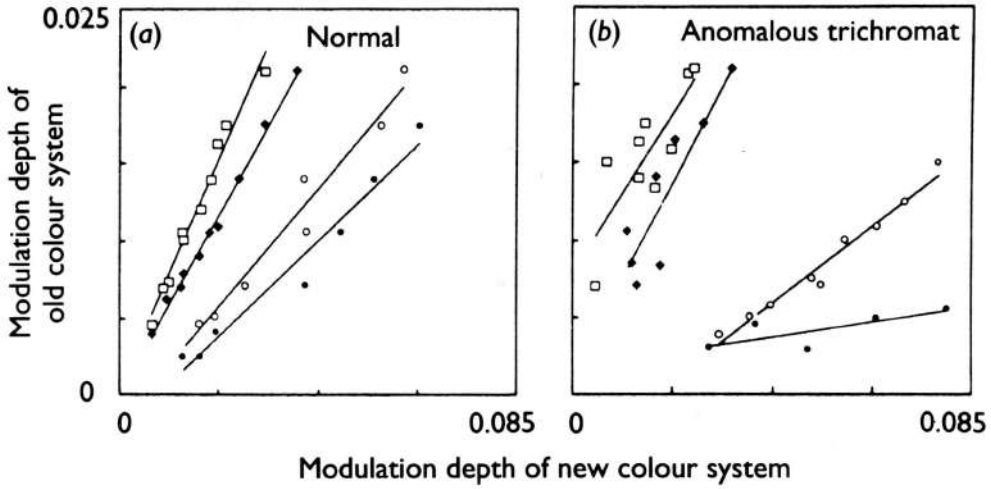

Figure 8 Results for (*a*) a normal observer and (*b*) an anomalous trichromat in a test that measures the relative salience of the signals of the phylogenetically old and new colour systems. Within each panel there are four sets of data points, corresponding to different separations of the elements in the array. Symbols as follows: (□) 0.144° of separation; (◆) 0.181° (○) 0.301°; (●) 0.602°. The ordinates of the graphs are expressed in units of the chromaticity diagram introduced by MacLeod and Boynton (*Journal of the Optical Society of America*, **69** (1979), 1183–6), a diagram whose axes correspond to the two subsystems discussed here. As the separation of elements increases, the older subsystem becomes the more salient, and this effect is exaggerated for the anomalous trichromat.

In other words, a given modulation of the newer system is balanced by a modest modulation of the old. Almost certainly, these results reflect the chromatic aberration of the eye and the sparse distribution of short-wave cones that I described earlier. They have implications for graphic designers and for those who program screens of information for computer displays: if graphic elements or stroke widths are small or are to be discriminated by the public at large viewing distances, then categories of information should not be differentiated by colours that differ mainly in the excitation of the short-wave cones.

Figure 8(*b*) shows analogous results for an anomalous trichromat and give us some insight into his private colour world. When the elements of the array are close together, the relative saliences of the old and new colour signals are similar to normal, but when the elements are well separated the old colour signal is relatively much the more salient and the

world of the anomalous trichromat is dominated by the warm–cold dimension.

The central representation of colour

The reader might suppose that the two subsystems I have been describing correspond to the traditional blue–yellow and red–green axes of colour space. There are four hues that to most people appear phenomenally pure or unmixed – red, yellow, green and blue – whereas other hues appear phenomenally to be mixtures: we feel we can see the red and the yellow in orange. Moreover, we never experience reddish greens or bluish yellows. On the basis of such phenomenological observations, the nineteenth-century physiologist Ewald Hering was led to postulate two corresponding pairs of antagonistic processes in the visual system, a yellow–blue process and a red–green one. Now, contrary to what one might read in textbooks, there is no physiological evidence in the visual system for cells that secrete the sensations of yellowness and blueness or redness and greenness. The two subsystems found in the retina and visual pathway simply do not correspond to Hering's two axes. A light that varied between pure yellow and pure blue, for example, would strongly modulate both subsystems, whereas a variation between violet and lime yellow is what is needed to stimulate exclusively the ancient colour system (see Figure 9).

But do the two subsystems remain independent at a cortical level? Michael Webster and I have been tackling this question by measuring the loss of saturation that occurs for some hues if the eye is adapted to a repetitive variation in colour. A reason for mentioning these experiments is that they illustrate one of the analytical tools available to experimental psychologists. By adapting to a repetitive stimulus of a carefully designed kind, we can seek selectively to adapt a particular mechanism that we suppose may be within the nervous system. In the present case we adapt the eye to variation along a particular axis in colour space (Figure 9). By afterwards probing sensitivity in this and other directions, we can estimate how narrowly turned are the central channels that represent colour and how many there are.

In our experiments, the subject views a computer-controlled colour

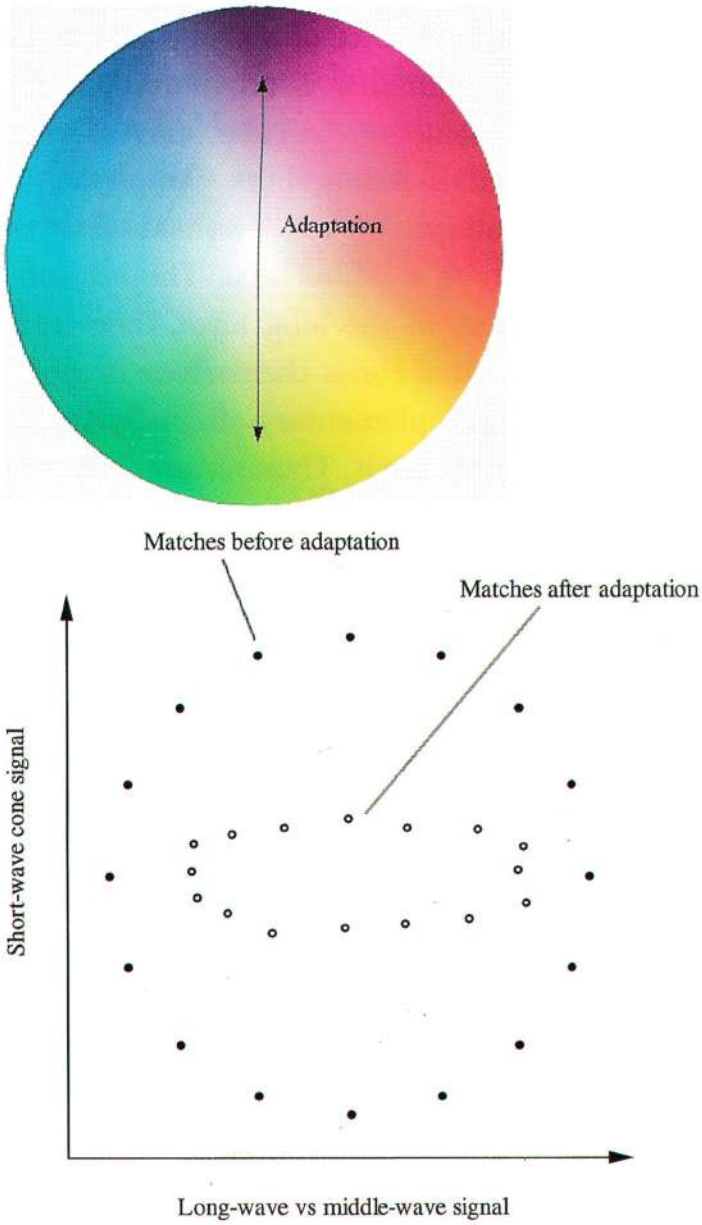

Figure 9 Probing central channels of colour vision by adaptation. The particular adaptation axis indicated above (from lime green to violet) is the one that would selectively adapt the ancient subsystem of colour vision. The subject was asked to adjust a reference colour falling on unadapted retina to match a test colour falling on adapted retina. In the graph below, coloured stimuli are plotted according to the signals they produce in the two subsystems, the vertical axis corresponding to the ancient subsystem and the horizontal axis to the newer one. The data points represent the appearances of colours before (●) and after (○) adaptation along a vertical axis: there is a selective loss of saturation along the same axis. The maximal loss of sensitivity is always along the adapting axis, whatever its direction (see text).

monitor. His or her task is to adjust the colour falling on an unadapted retinal area so as to match the appearance of a probe colour that falls on an adapted area. The adapting stimulus is an illuminated patch that varies rhythmically in colour along a particular axis of colour space, changing, for example, between a red and its complementary bluish green once per second. For all adapting axes, we ensure that the average stimulus over time is equivalent to a white light. The advantage of this kind of adapting stimulus is that, in the long term, there should be no differential adaptation of the retinal cones themselves; indeed we do not observe the common-or-garden complementary after-effects that one sees after simply staring at a coloured patch. Thus we are able to probe the neural representation of colour at a more central level.

In the graph of Figure 9, the outer solid circles represent the subject's match before adaptation. Suppose now we view an adapting field that varies between violet and yellow-green and thus exercises only the ancient subsystem. The open circles in the figure represent the match after adaptation. Colours that lie along the adapting axis then look very desaturated and plot closer to white, at the centre of the diagram, whereas colours on the orthogonal axis are relatively little altered in appearance. A similarly selective result is seen if the adaptation is along a horizontal axis from red to bluish green: now the greatest loss of saturation is found for stimuli along the horizontal axis. But if these axes were cardinal, if they corresponded to two classes of cortical neuron that between them represented all colours, then what should we expect if we adapted along a 45° axis in this space? If colour were represented only by the two cardinal axes, there should now be an equal loss of saturation for all colours. But in fact we again find that the saturation losses are largest along the axis of adaptation, and this is true for whatever direction in colour space we try.

This suggests that at the cortical level, there are neural channels tuned for many different directions in colour space. We do observe some enhanced selectivity for two axes, but these axes correspond to the ancient and modern subsystems identified physiologically in the visual pathway – and not to the putative red–green and yellow–blue processes of Hering. The special phenomenal status of the four pure hues is perhaps the chief unsolved mystery of colour science.

Colour constancy

I must deal finally with the relativity of our colour perception. The great geometer Gaspard Monge has attracted from his countrymen the supreme tribute of an eponymous Metro station, but what to me is a mark of his distinction is that he held high office, and retained his head,

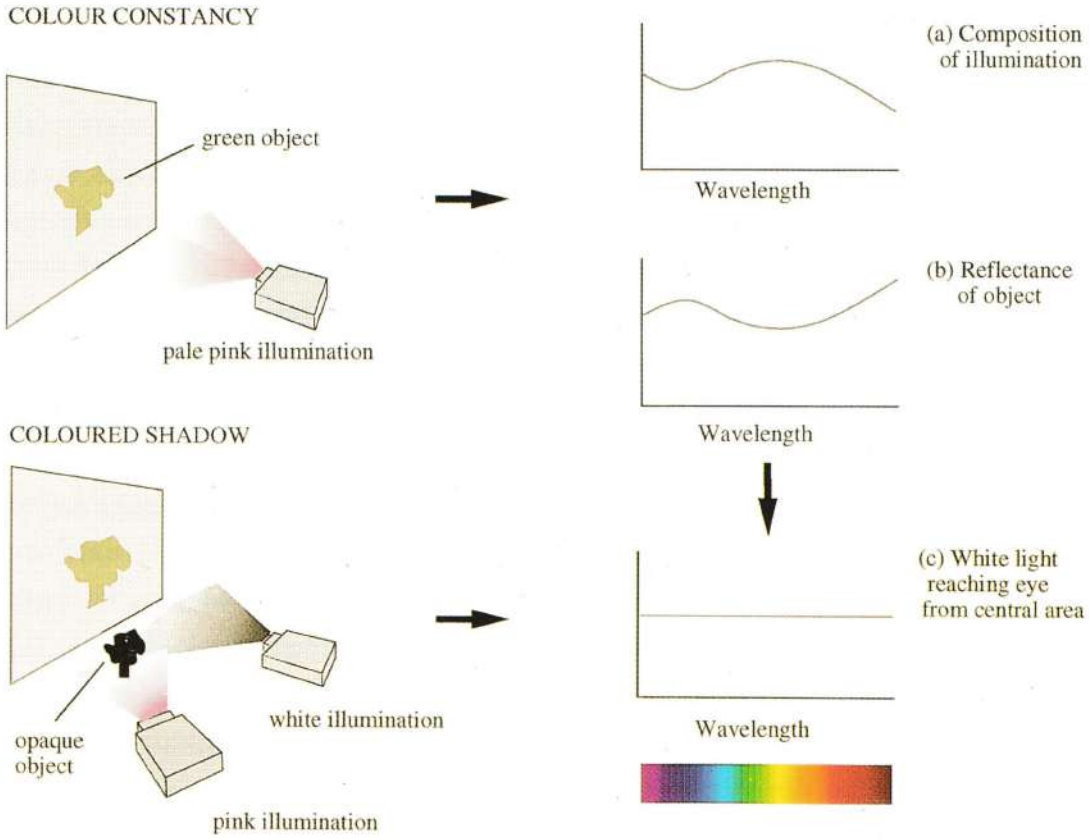

Figure 10 A coloured shadow as a demonstration of colour constancy. At the bottom left a white screen is illuminated by light from two projectors, one giving pink light and one giving white light. An opaque card in the pink beam casts a shadow on the screen. This area therefore reflects only white light to the eye (graph *c*), but it appears a quite vivid, bluish green to the audience. Above is shown how a truly blue-green object would present the same stimulation to the eye if it were presented on a white ground in a pale pink illumination (top left). If we multiply the strength of the illumination at different wavelengths (graph *a*) by the proportion of light reflected from the object at each wavelength (graph *b*), then we derive (*c*) the relative intensity of different wavelengths reaching the eye from the central area. Graph (*c*) is identical in the case of 'colour constancy' and in the case of the 'coloured shadow'. It is also the case – and this is crucial – that the light reaching the eye from the surround is identical in the two cases.

in every administration from the *ancien régime* to the First Empire. On the eve of the Revolution, in May 1789, he gave to the Académie Royale des Sciences the most brilliant lecture that has ever been given on colour perception. We know, from an eyewitness account, the experiment with which he began: on the wall of a house that faced the windows of the Academy, he had fixed a sheet of red paper. He invited his fellow *académiciens* to look through a red glass and consider the colour of the red paper. The result was as counterintuitive in 1789 as it is today. We might expect that the red paper seen through red glass would have looked a vivid scarlet. In fact, it appeared whitish, and much less strongly coloured than when viewed with the naked eye. The experiment is today easy to repeat using a piece of red gelatine, and – as Monge himself remarked – the effect is most striking when the visual scene is a complex one, containing a variety of surfaces of different colours as well as the red target. A busy carpark serves well.

Monge realised that this was not an isolated illusion. He recognised that it shared an explanation with another phenomenon, which was already ancient in 1789, the phenomenon of coloured shadows. We can illustrate this effect by illuminating a screen with two projectors, one casting white light on the screen, the other, pink light – so that the surface appears homogeneously illuminated with very pale pink (Figure 10). If now we introduce a small opaque silhouette into the pink beam, so that it casts a shadow on the screen, then the shadow will appear a rather vivid green. If the demonstration is well set up, the green of the shadow will seem much more saturated than the pale pink of the rest of the screen. And yet – here is the mystery – the area of the shadow is physically illuminated only by white light.

Monge explained, with the frightening clarity that can be achieved only in the French language, that coloured shadows and the paradox of the red filter are not isolated illusions. Rather they reveal the operation of a mechanism that serves us well every instant of the day. This is the process of colour constancy, the process that allows us to perceive the permanent surface colours of objects independently of the colour of the illumination. A piece of white paper stubbornly looks white, whether we examine it in the yellowish light of indoor tungsten or in the bluish cast of northern daylight. What our visual system is built to recognise is per-

manent properties of objects, their spectral reflectances, their permanent tendency to reflect one wavelength more or less than another. What we are not built to recognise is the spectral flux, the balance of wavelengths in the light reaching our retina from an object on a particular occasion. For the latter, the balance of wavelengths reaching our eye from the object, depends both on the surface properties of the object and on the colour of the illumination.

But how does the visual system separate the spectral composition of the illuminant from the spectral reflection of the object? Monge supposed that our visual system estimates the illuminant from the average spectral flux in the surrounding field. In the demonstration of coloured shadows, our visual system very reasonably takes the illuminant to be a pale pink. The shadow is an object that reflects proportionately less red light than does the rest of the field. An object that behaves this way in pinkish illumination must be a green object.

We can use Figure 10 to elaborate the relationship between the illusion we call coloured shadows and the normal process we call colour constancy. In the case of the coloured shadow demonstration, the light reaching our eye from the shadowed area is white (as shown on the right) and the surround is a pale pink. In the equivalent case of colour constancy, we have an actually green object in pale pink illumination. If we multiply the spectral composition of the illumination by the spectral reflectance of the green object, we again have spectrally flat light coming from the object. The visual system is very intelligent, but it is not so intelligent as to distinguish between two projectors and one; these two stimulus arrays are therefore equivalent.

So there is no fixed relationship between the spectral composition of light and the hue that we perceive at a local point in the scene. This is the conclusion drawn with such prescient clarity by Gaspard Monge in 1789, when he wrote:

> So the judgements that we hold about the colours of objects seem not to depend uniquely on the absolute nature of the rays of light that paint the picture of the objects on the retina; our judgements can be changed by the surroundings, and it is probable that we are influenced more by the ratio of some of the properties of the light rays than by the properties themselves, considered in an absolute manner.

In other words, we judge colours by the company they keep.

Further reading

Cronly-Dillon, J. R. (ed.), *Vision and Visual Dysfunction*, vols. 2, 6 and 7, London: Macmillan, 1991.

Furumoto, L., 'Joining separate spheres – Christine Ladd-Franklin, woman-scientist (1847–1930)', *American Psychologist*, **47** (1992), 175–82.

Jacobs, G. H., 'The distribution and nature of colour vision among the mammals', *Biological Reviews*, **68** (1993), 413–71.

Jordan, G., and Mollon, J. D., 'A study of women heterozygous for colour deficiencies', *Vision Research*, **33** (1993), 1495–508.

Mollon, J. D., and Bowmaker, J. K., 'The spatial arrangement of cones in the primate fovea', *Nature*, **360** (1992), 677–9.

Monge, G., 'Mémoire sur quelques phénomènes de la vision', *Annales de Chimie*, **3** (1789), 131–47.

Nathans, J., 'The genes for color vision', *Scientific American*, **260**(2) (1989), 29–35.

Neitz, M., Neitz, J., and Jacobs, G. H., 'Spectral tuning of pigments underlying red–green color vision', *Science*, **252** (1991), 971–4.

Pokorny, J., Smith, V. C., Verriest, G., and Pinckers, A. J. L. G., *Congenital and Aquired Color Vision Defects*, New York: Grune & Stratton, 1979.

Webster, M. A., and Mollon, J. D., 'Changes in colour appearance following post-receptoral adaptation', *Nature*, **349** (1991), 235–8.

Winderickx, J., Lindsey, D. T., Sanocki, E., Teller, D. Y., Motulsky, A. G., and Deeb, S. S., 'Polymorphism in red photopigment underlies variation in colour matching', *Nature*, **356** (1992), 431–3.