

## “THO’ SHE KNEEL’D IN THAT PLACE WHERE THEY GREW...”

### THE USES AND ORIGINS OF PRIMATE COLOUR VISION

BY J. D. MOLLON

*Department of Experimental Psychology, University of Cambridge, Downing St,  
Cambridge CB2 3EB*

#### Summary

The disabilities experienced by colour-blind people show us the biological advantages of colour vision in detecting targets, in segregating the visual field and in identifying particular objects or states. Human dichromats have especial difficulty in detecting coloured fruit against dappled foliage that varies randomly in luminosity; it is suggested that yellow and orange tropical fruits have co-evolved with the trichromatic colour vision of Old World monkeys. It is argued that the colour vision of man and of the Old World monkeys depends on two subsystems that remain parallel and independent at early stages of the visual pathway. The primordial subsystem, which is shared with most mammals, depends on a comparison of the rates of quantum catch in the short- and middle-wave cones; this system exists almost exclusively for colour vision, although the chromatic signals carry with them a local sign that allows them to sustain several of the functions of spatiochromatic vision. The second subsystem arose from the phylogenetically recent duplication of a gene on the X-chromosome, and depends on a comparison of the rates of quantum catch in the long- and middle-wave receptors. At the early stages of the visual pathway, this chromatic information is carried by a channel that is also sensitive to spatial contrast. The New World monkeys have taken a different route to trichromacy: in species that are basically dichromatic, heterozygous females gain trichromacy as a result of X-chromosome inactivation, which ensures that different photopigments are expressed in two subsets of retinal photoreceptor.

#### Introduction

In his *Uncommon Observations About Vitiated Sight*, a collection of case reports published in 1688, Robert Boyle reproduces his notes on a gentlewoman whom he had examined some years earlier (Fig. 1). During an unidentified illness, apparently treated by fierce vesication, this unhappy maid had transiently gone blind. Visual function had gradually returned, and Boyle satisfied himself that she could read from a book and had good acuity.

‘...having pointed with my Finger at a part of the Margent, near which

Key words: colour vision, polymorphism, primates.

SOME  
UNCOMMON  
OBSERVATIONS  
ABOUT  
VITIATED  
SIGHT.



LONDON:  
Printed for J. Taylor, at the Ship in  
St. Paul's Church-Yard, 1688.

( 265 )

ces, that I now with had not been omitted.

The Gentlewoman I saw to day, seems to be about 18 or twenty years old, and is of a fine Complexion, accompanied with good Features. Looking into her Eyes, which are Gray, I could not discern any thing that was unusual or amiss; tho' her Eye-lids were somewhat Red, whether from Heat, or which seemed more likely, from her precedent Weeping. During the very little time that the Company allowed me to speak with her, the Questions I propos'd to her were answered to this Effect.

That about five years ago, having been upon a certain Occasion immoderately tormented with Blisters, applied to her Neck and other Parts, she was quit deprived of her sight.

Ibat

( 257 )

look'd attentively upon it, but told me, that to her it did not seem Red, but of another Colour, which one would guess by her Description to be a Dark or Dirty one: and the Tufts of Silk that were finely Colour'd, she took in her Hand, and told me they seem'd to be a Light-colour, but could not tell me which; only she compar'd it to the Colour of the Silken Stuff of the Lac'd Peri-coat of a Lady that brought her to me; and indeed the Blews were very much alike. And when I ask'd her, whether in the Evenings, when she went abroad to walk in the Fields, which she much delighted to do, the Meadows did not appear to her Cloathed in Green? she told me they did not, but seem'd to be of an odd Darkish Colour; and added, that when she had a mind to gather Violets, tho' she kneel'd in that Place where they grew, she was not able to distinguish them by the Colour from the neighbouring Grass, but only by the

( 268 )

the Shape, or by feeling them. And the Lady that was with her, took thence occasion to tell me, that when she looks upon a Turkey Carpet, she cannot distinguish the Colours, unless of those parts that are White or Black. I ask'd the Lady whether she were not troubled with Female Obstructions? To which she Answer'd me, she was not now, but that formerly she had been much subject to them, having been obstinately troubled with the Green-sickness.

Fig. 1. Title page and excerpts from Robert Boyle's *Vitiated Sight* of 1688. (Note: the original pagination is in error.)

there was the part of a very little Speck, that might almost be covered with the point of a Pin; she not only readily enough found it out, but shewed me at some distance off another Speck, that was yet more Minute...’.

But what was most singular and strange for Boyle was that she was left with a permanent loss of colour vision: ‘...she can distinguish some Colours, as Black and White, but is not able to distinguish others, especially Red and Green...’.

Boyle’s early case of acquired achromatopsia, and similar cases described in modern times (e.g. Meadows, 1974; Damasio *et al.* 1980; Mollon *et al.* 1980), bear on one of the central themes of this symposium, the question of whether there are submodalities within each of the major sensory modalities: if colour vision can be disproportionately impaired, while processing of spatial detail remains relatively intact, then we must suspect that there is some independence between the analysis of colour and the analyses of other attributes of the retinal image. The issue is less new than we might suppose: as early as 1869, in discussing the recovery of spatial vision before chromatic vision in a patient recovering from optic neuritis, J. J. Chisholm asked explicitly ‘whether there are special nerve fibres, for the recognition of special colours, independent of those used in the clear definition of objects’.

But a second interest of Boyle’s gentlewoman lies in the practical expression of her disability.

‘...that when she had a mind to gather Violets, tho’ she kneeled in that Place where they grew, she was not able to distinguish them by the Colour from the neighbouring Grass, but only by the shape, or by feeling them.’

Her disability reminds us that we can best understand primate colour vision, and its evolution, by considering the price of not enjoying colour vision: what are the advantages that we should then forego?

### **The advantages of colour vision**

#### *Detection of targets against dappled backgrounds*

In the current literature it is commonly suggested that colour vision serves particularly to detect edges between ‘equiluminant’ surfaces, surfaces of the same luminance but different chromaticity. Such a view was systematically expressed, for example, by Gouras & Eggers (1983). But, in fact, it is rare in the natural world for one surface to lie in front of another in such a way that both have the same reflectance, both lie at the same angle to the incident illumination, and the nearer throws no shadow on the farther.

And if Nature finds it hard to generate equiluminant edges, so does Art: in the late nineteenth century, when the first ‘pseudoisochromatic’ plates were constructed for detecting colour-deficient observers, attempts were made to print targets of one colour on fields of different colour in such a way that dichromats could not detect them. The enterprise was not long pursued, since it was hardly possible for the printer so exactly to match figure and field in lightness and texture, and so completely to eliminate edge artefacts, as to deceive the average

dichromat. But in the early pseudoisochromatic plates of Stilling (1883), like the later ones of Ishihara, the problem was neatly sidestepped by two ploys: first, the target was formed from discrete patches, embedded in a background array of similar patches; and second, the lightnesses of the individual patches were varied randomly rather than being equated (see Fig. 2).

Such pseudoisochromatic plates, which have proved so successful an artifice for detecting colour-deficient observers, nicely show us the tests that defeat the colour-deficient in the natural world. The colour-deficient observer is not often confronted by a homogeneous target of one chromaticity that lies on an equiluminant surface of different chromaticity. Rather, his difficulty arises when the background is dappled and brindled, when, that is, luminosity is varying randomly. Such a variation in luminosity can arise because the illuminant is interrupted by foliage; or it can arise because the background consists of component surfaces that lie at varying angles to the illuminant or (as in the case of new and old leaves) that themselves vary in reflectance. Boyle's gentlewoman is poor at spotting violets among grass: she cannot use lightness to solve the task, and is forced to use shape, but under conditions such that the presence of similar, masking, contours slows down a search on the basis of shape. In the early reports of colour blindness, a recurrent theme is difficulty in searching for natural targets against foliage, vegetation or other visually busy backgrounds. The celebrated shoemaker Harris, described by Huddart in the *Philosophical Transactions* for 1777, had noticed as a child that other children could detect cherries on a tree at a greater distance than he could.

'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.'

Similarly, the observer described by Nicholl (1818) reports:

'...the fruit on trees when red, I cannot distinguish from the leaves, unless when I am near it, and then from the difference of shape rather than colour'.

The protanope of Colquhoun (1829) had been perverse enough to become an orchardist after colour blindness had stymied his earlier career as a weaver.

'He cannot discern, even in a loaded bush, the existence of red gooseberries [sic] among the leaves, until he has almost approached so near as to be able to take hold of the branch. Rosy apples on a tree, which may be discovered by ordinary eyes, at a distance of from thirty to forty yards at least, are lost to his sense, until he has come within ten or twenty yards of the tree, when he can trace out the fruit by its form...'

The colour-deficient brassfounder of Edinburgh, described by Combe (1824), was a keen shot, but could not 'discover game upon the ground, from the faintness of his perception of colours'.

'...last year, when a large covey of partridges rose within ten or twelve yards of him, the back ground being a field of Swedish turnips, he could not perceive a single bird.'



Such complaints are familiar in the early literature on daltonism, although few authors recapture the insight of Huddart's telling phrase '*...if they were not enveloped in other things, as...cherries among the leaves.*' But, however commonplace the complaint might seem, we should listen to it. It tells us the disability of the daltonic observer; and the complement of his disability must be the advantage of the normal trichromat. *A primary advantage of colour vision is that it allows us to detect targets against dappled or variegated backgrounds, where lightness is varying randomly.*

### *Perceptual segregation by colour*

The pseudoisochromatic plates of Stilling or of Ishihara show us a second way in which colour assists our perception: the shared hue of different elements in a visual array may serve as a basis for perceptual segregation and so allow the patient to identify a figure which is made up from discrete patches that vary in lightness.

Segregation of the visual field into elements that belong together is an important preliminary to the actual recognition of objects. This segregation corresponds to the process that the Gestalt psychologists called 'figure-ground differentiation' or 'perceptual organization'. Gestalt writers recognized that many features of the visual image – brightness, colour, local form, local texture, binocular disparity, direction of movement – can each serve as a basis for perceptual organization (Wertheimer, 1923; Koffka, 1935). Barlow (1981) used the term 'linking features' to describe stimulus attributes that can sustain perceptual segregation.

Livingstone & Hubel (1988) make the remarkable claim that 'only luminance contrast, and not color differences, is used to link parts together'. This claim should be understood exclusively in terms of the particular matter with which they are concerned, the incoherence of the visual field when edges are defined only by chromatic differences: under such conditions of 'equiluminance', it becomes difficult, for example, to relate different parts of colinear edges, and we lose the coordination that would otherwise impose a uniform interpretation on apparent movements in different parts of the field. But perceptual failures of this kind may have a special explanation: under the unusual conditions of equiluminance the long-wave receptors of the retina are signalling spatial and temporal transients of opposite, contradictory, sign from those signalled by the middle-wave receptors (see Mollon, 1980, 1989). The very existence of pseudoisochromatic plates shows that Livingstone & Hubel's statement cannot be generally true: normal observers can perfectly well use colour to link different elements in the visual field and extract a figure. Notice, in the case of the Stilling or Ishihara plates (Fig. 2), that this is done under conditions where lightness is varying randomly – conditions that are very much more common in the natural world than are conditions where lightness remains constant across the field.

When we recognize a digit in the Ishihara plates, the contours of individual patches are detected by their luminance contrast, but the digit is discovered only by colour. The cortex must be able to collect together patches of similar chromaticity while yet preserving the local signs – the spatial addresses – that

belong with these patches; and must be able to deliver this subset of spatial addresses as the input to a form-recognition system. We do not know in detail how this is done, although it is plausible to suppose (with Barlow, 1981) that the correlation of patches of similar chromaticity is achieved in the prestriate region in which colour is emphasized (Zeki, 1978) – V4 in the rhesus monkey and perhaps the region of the lingual and fusiform gyri in man (Meadows, 1974); and it is also plausible to suppose that the set of spatial addresses is ultimately delivered to a common form-recognition system, which is located in the inferotemporal lobe and is able to accept similar sets of spatial addresses defined by other linking features such as movement, texture, depth and orientation. But whatever cortical machinery collects common chromaticities while preserving spatial addresses, the important point for our present purpose is that such machinery must evolve before an animal can use colour for the particular purpose of perceptual segregation.

Is there any price to be paid for the advantage of being able to segregate the field on the basis of colour? Here again the Ishihara plates point to the answer. There exist plates (e.g. plates 18–21 of the 10th edition) where a digit is readily visible to the dichromat, on the basis of differences in the short-wave component, but is masked for many normals, owing to a stronger, rival, organization suggested by red–green variations. Morgan *et al.* (1989) prepared a visual display in which one quadrant of the array differed from the rest in the size or orientation of texture elements; they found that the textural difference could be masked for normals if a random variation in colour was introduced, but that dichromats performed as well as they did in a monochrome condition. A fundamental limitation of the visual system is that it cannot concurrently entertain different perceptual organizations; and this, in itself, suggests that figural organizations based on colour, texture, orientation and other attributes are delivered to a common pattern-recognition system.

#### *Identification by colour*

Finally, colour may serve to identify, that is (a) it may help us assign an already-segregated object to a given category or (b) it may indicate what lies beneath a given surface.

Colour may be one of the cues that monkeys use to identify the species of particular trees or particular plants. Certainly, it must be one of the cues that catarrhine (Old World) monkeys use to identify conspecifics, the sex of conspecifics, and the sexual state of conspecifics. In many members of the genus *Cercopithecus*, the slate blue scrotum of the male contrasts with surrounding yellow fur; and often this contrast is echoed in facial markings, most notably in the moustached guenon (*Cercopithecus cephus*) whose vivid blue face is surrounded by a yellow ruff. Such colourings in *Cercopithecus* are often accompanied by species-specific ‘flagging’ movements and may serve to maintain species isolation in habitats where several species of guenon overlap (Kingdon, 1980, 1988). Other striking examples of the use of colour as sexual signals are offered by the bright red

ischial callosities of the female baboon (*Papio papio*) and the blue, violet and scarlet patterns of the genitalia and muzzle of the male mandrill.

Colour offers a means to discover remotely the structure underlying a surface (Katz, 1935). In the case of fructivorous primates, one of the most important functions of trichromatic colour vision must be to judge the state of ripeness of fruit from the external appearance (Gautier-Hion, 1988). Similarly, variations in the colour of ground-covering vegetation can reveal at a distance the presence of water. In this category of the use of colour, we might also include judgements of complexion: human observers have a finely developed (though little studied) capacity to use skin colour to estimate the health, or emotional state, of conspecifics, and this ability may have biological advantage in the selection of sexual partners or in the care of infants.

In so far as an animal uses colour to identify – in so far as it uses colour vision to recognize permanent attributes of an object or to make absolute judgements about the spectral reflectances of surfaces – we should expect it to exhibit colour constancy: it ought to be able to recognize the spectral reflectance of a surface independently of the spectral composition of the illuminant (see e.g. Katz, 1935; Land, 1983). Of the three uses of colour vision distinguished here, only the third necessarily requires colour constancy; but the distinction between the three functions is not absolute and there will be many cases, for example, where an animal is searching a complex array not merely for a salient stimulus but for a stimulus of a particular spectral reflectance.

For an alternative approach to the uses of colour vision, the discussion by Jacobs (1981) is recommended.

### Trichromacy and its evolution

#### *The photoreceptors*

In man, the three classes of retinal cone exhibit overlapping spectral sensitivities, with peak sensitivities lying close to 420, 530 and 560 nm (Fig. 3). Estimates of these sensitivities have now been obtained by psychophysics (Boynton, 1979), by microspectrophotometry of the outer segments of individual cones (Darnall *et al.* 1983), and by electrophysiological measurements of isolated receptors that have been sucked into micropipettes (Schnapf *et al.* 1987); and the different estimates agree to a first approximation. Notice the asymmetry of the spectral positions of the receptors: there is an interval of 100 nm between the short-wave and middle-wave receptors, but an interval of only 30 nm between the middle-wave and long-wave receptors.

The response of any individual receptor obeys the Principle of Univariance: its electrical response varies only with the rate at which photons are captured by the outer segment, and all that changes with wavelength is the probability that any given photon will be absorbed. Thus, if the visual system is to learn about wavelength, independently of radiance, it must not only have receptors with



Fig. 2. Test patterns from the pseudoisochromatic plates of Stilling. Note that the figure and field are made up of discrete patches, varying in lightness and each having its own contour.

Fig. 3. The absorbance spectra of the human cone pigments. These curves are replotted from Table 2 of Dartnall *et al.* (1983) and were obtained by microspectrophotometric measurements of individual photoreceptors from human eyes removed in cases of ocular cancer. The ordinate represents the (normalized) logarithm of the factor by which monochromatic light is attenuated as it passes through the outer segment of an isolated receptor.

different sensitivities but also have the neural machinery to measure the relative rates of photon capture in different classes of receptor (Rushton, 1972).

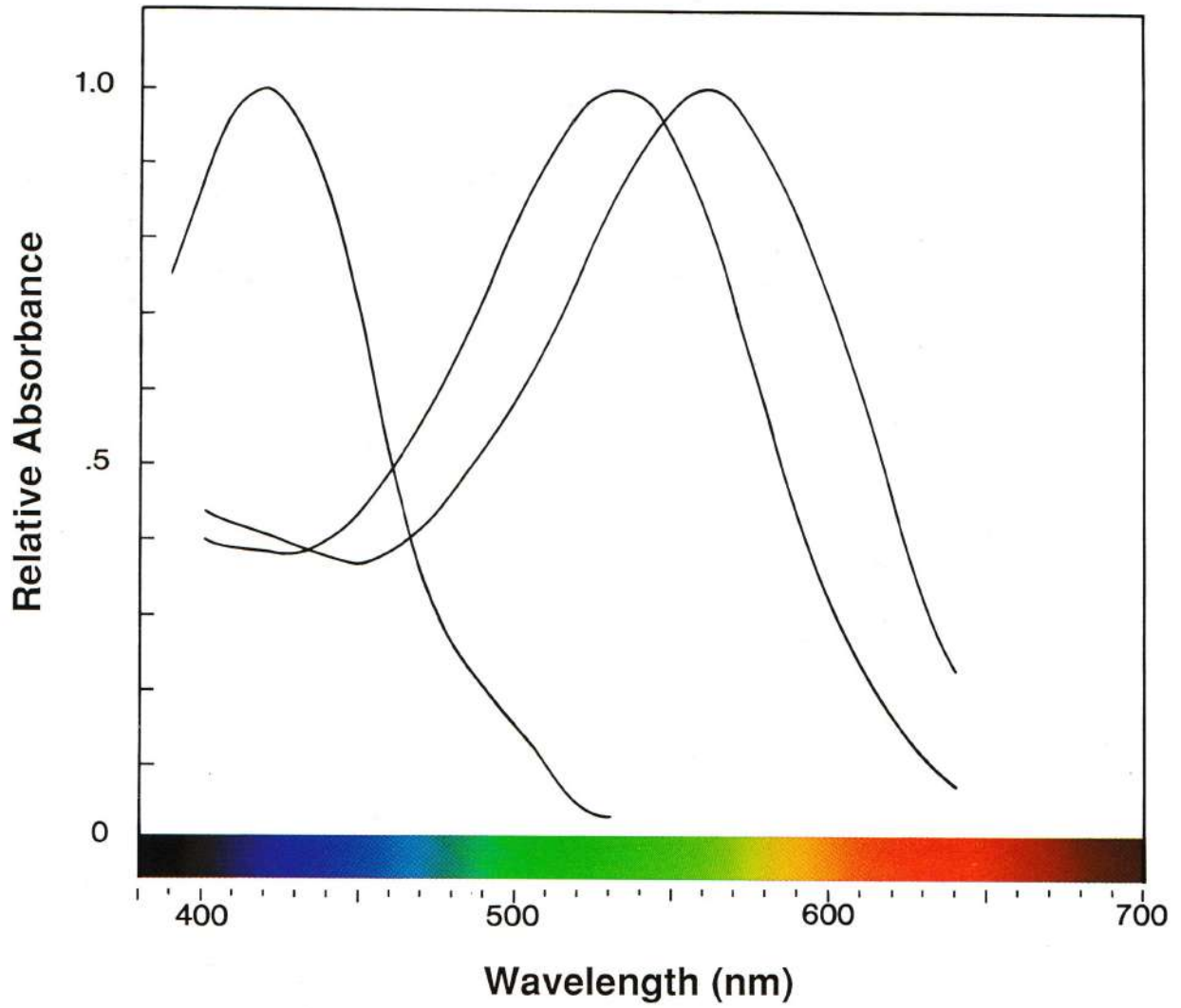
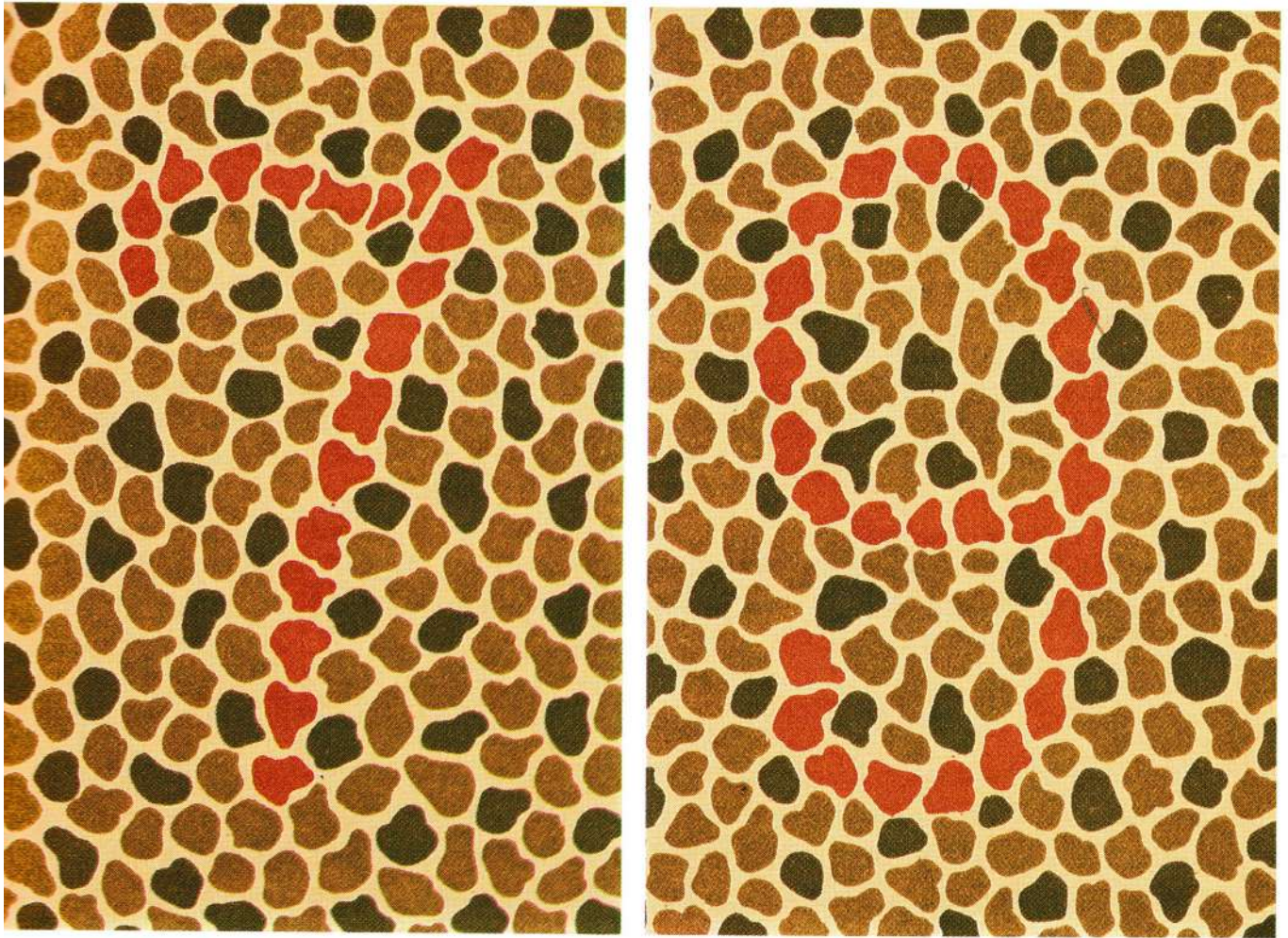
Textbooks often represent the three types of human cone as equal elements in a trichromatic scheme, as if they evolved all at once and for a single purpose. All the evidence suggests, however, that they evolved at different times and for different reasons; and our own colour vision seems to depend on two, relatively independent, subsystems – a phylogenetically recent subsystem overlaid on a much more ancient subsystem (Mollon & Jordan, 1989).

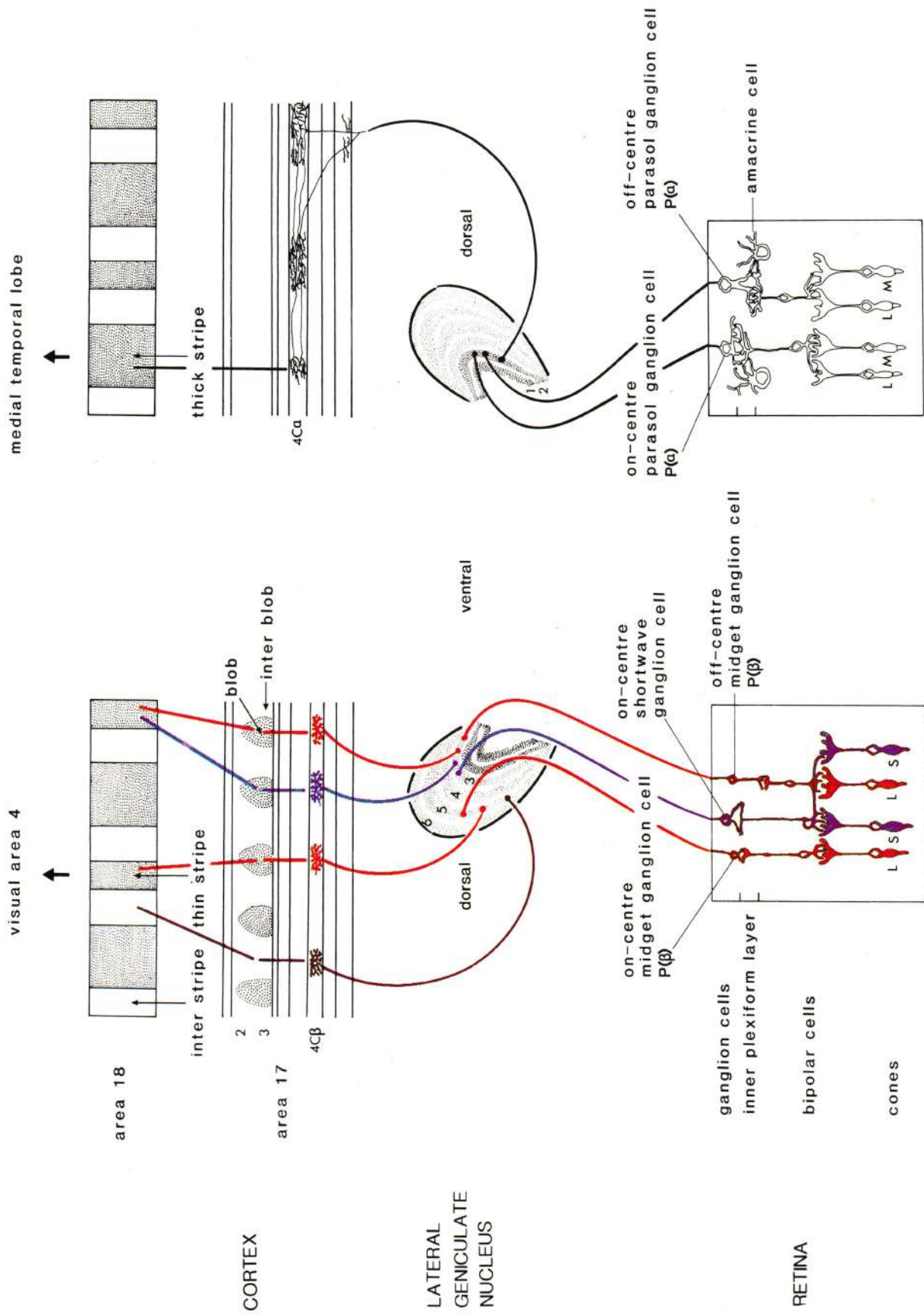
#### *The primordial subsystem of mammalian colour vision*

For the main business of vision, for the detection of contours, movement and flicker, most mammals depend on a single class of cone, which has its peak sensitivity near the peak of the solar spectrum, in the range 510–570 nm (Jacobs, 1981). A long time ago, a few short-wave (violet-sensitive) cones were sparsely sprinkled amongst the matrix of middle-wave cones. These short-wave cones, with their wavelength of peak sensitivity well removed from that of the middle-wave cones, provided mammals with a basic, dichromatic, system of colour vision: by comparing the quantum catch of the short-wave cones with that of the middle-wave cones, the visual system is able to estimate the sign and the slope of the change in spectral flux from one end of the spectrum to another. If the spectral distribution of the light reaching us from an object is itself considered as a waveform (Barlow, 1982), then this primordial colour-vision system seems designed to extract the lowest Fourier component of the waveform (Mollon *et al.* 1989). Subjectively, it is this primordial subsystem that divides our colour sensations into warm, cool and neutral (McDougall, 1901). And it is this subsystem that survives in the common forms of human dichromacy.

The antiquity of the primordial subsystem is suggested by the molecular genetic results of Nathans *et al.* (1986*b*), who have sequenced the genes that code for the protein moieties of the four human photopigments. The amino-acid sequence for the short-wave cone pigment is as different from those for the long- and middle-wave pigments as it is from the sequence for the human rod pigment, rhodopsin: in a pairwise comparison, the short-wave pigment shares 42 % of its amino acids with rhodopsin, 43 % with the long-wave cone pigment and 44 % with the middle-wave cone pigment. Nathans and his colleagues use the difference in sequence between bovine and human rhodopsin to estimate the rate at which photopigments diverge,







MAGNOCELLULAR PATHWAY

PARVOCELLULAR PATHWAY

Fig. 4



Fig. 4. The parallel channels of the geniculostriate pathway of the primate visual system. To the right is represented the magnocellular system, which carries little chromatic information: the large  $P\alpha$  cells of the retina project to the magnogeniculate nucleus and thence to layer  $4C\alpha$  of the visual cortex and the 'thick stripes' of area 18, which stain strongly for cytochrome oxidase. A second parallelism is represented within the parvocellular pathway on the left; and here the parallelism is between a channel that carries almost purely chromatic information (the primordial subsystem of colour vision) and a channel that carries information about both colour and spatial contrast (the second subsystem of colour vision). The two subsystems, represented by blue and red, respectively, appear to remain independent, at least as far as the cytochrome oxidase 'blobs' of area 17 of the cortex.

---

and they conclude that the short-wave pigment diverged from the middle-wave cone pigment (or perhaps indeed from an ancestral rod pigment) more than 500 million years ago.

#### *The primordial subsystem considered as a purely chromatic channel*

It would be tempting to say that the short-wave cones have been given to us only for colour vision, and that the pathway which carries their signals is the nearest we come to a channel that conveys purely chromatic information. There is certainly good reason why the short-wave cones should not be used for the discrimination of spatial detail: the short-wave component of the retinal image is almost always degraded, since bluish, non-directional skylight dilutes the shadows of the natural world and since the eye itself is subject to chromatic aberration and usually chooses to focus optimally for yellow light. The short-wave cones were therefore added sparingly to the retinal matrix, and they remain rare: they constitute only a few per cent of all cones in every primate species where cones have been identified directly by microspectrophotometry (e.g. Bowmaker *et al.* 1987; Hárosi, 1982; Dartnall *et al.* 1983; Mollon *et al.* 1984). At a post-receptoral level, little sensitivity to spatial contrast is exhibited by those neurones that draw signals from the short-wave cones (Derrington *et al.* 1984): such cells are *chromatically* opponent but tend not to draw their opposed inputs from clearly distinct, spatially antagonistic, regions in the way that is usual for other cells in the retina and lateral geniculate nucleus.

It would, in fact, be careless to say that the information provided by the short-wave cones is *purely* chromatic. Their signals must carry with them enough of a local sign that the chromatic information can be associated with the corresponding object. And the local signs are adequate to support perceptual segregation (see above), as is demonstrated by those specialized pseudoisochromatic plates that bear figures visible to the normal trichromat but invisible to tritanopes (the rare type of dichromats who lack the short-wave cones): examples of such plates are those of Willmer (see Stiles, 1952) and of Farnsworth (see Kalmus, 1955). Moreover, if the short-wave cone signals are isolated by presenting a violet target on an intense long-wave adapting field, then the contours formed directly by short-wave signals can support the perception of gratings of low spatial frequency (e.g.

Brindley, 1954); such gratings can give an orientationally selective adaptation (Stromeyer *et al.* 1980) and can provide an input to a motion-detecting mechanism that also draws inputs from the long-wave cones (Stromeyer & Lee, 1987). M. A. Webster, K. K. De Valois and E. Switkes (in preparation) isolated the short-wave system by forming an equiluminant low-frequency grating of two colours that are confused by a tritanope; they found that the discrimination of orientation and spatial frequency was as good as it was for an equiluminant grating that was visible by virtue of variations in the ratio of the signals of the long-wave and middle-wave cones.

Nevertheless, in the natural world, the short-wave cones must play little role in the detection of sharp edges and the discrimination of spatial detail. Using a checkerboard pattern, Stiles (1949) first demonstrated the low values that are obtained for foveal acuity when vision depends only on the signals of the short-wave cones. And later measurements of spatial contrast sensitivity suggest that the short-wave receptors cannot sustain a resolution of much greater than 10 cycles per degree (for references, see Mollon, 1982, Table 1); the maximal sensitivity for vision with these cones lies close to 1 cycle per degree. Tansley & Boynton (1976) required subjects to adjust the relative luminances of two hemifields of different chromaticity until the edge between them was minimally distinct; the subject then rated the salience of the residual edge on an eight-point scale. When the two half-fields differed only in the degree to which they excited the short-wave cones, when, that is, they lay on the same tritan confusion line in colour space, observers gave a rating of zero to the distinctness of the border between the two fields: the fields 'melted' into one another. And the rated distinctness of other (non-tritan) borders could be predicted simply by how much the two half-fields differed in the ratio with which they excited the long- and middle-wave cones (see also Thoma & Scheibner, 1980). In her classical description of 'hard' and 'soft' colours, Liebmann (1927) described how a blue figure perceptually merged (*verschwimmt*) with an equiluminant green ground; and in retrospect we may suspect that these two colours were ones that came close to lying along the same tritan confusion line. In the most recent literature on equiluminance the Liebmann effect has sometimes been cited, but often no distinction is made between the modest effects obtained when long- and middle-wave cones are modulated and the much more spectacular weakening of contours and of figural organization when the target and field differ only along a tritan confusion line, i.e. when they differ only in the degree to which they excite the primordial subsystem of colour vision.

#### *The morphological basis of the primordial subsystem*

There are growing hints that the primordial subsystem of colour vision has a morphologically distinct basis in the primate visual system (indicated in blue in Fig. 4). Thus Mariani (1984) has described a special type of bipolar cell, which resembles the common 'invaginating midget' bipolar but which makes contact with two well-separated cone pedicles; he suggests that these bipolars are exclusively in contact with short-wave cones. Using micropipettes filled with the dye Procion



Yellow, De Monasterio (1979) made intracellular recordings from (and stained) three ganglion cells that showed a maximal sensitivity for 440-nm increments in the presence of a long-wave field: the cell bodies of these short-wave ganglion cells were clearly larger than those of the standard midget ganglion cell that draws opposed inputs from the long- and middle-wave receptors. [Rodieck (1988) has drawn attention to a discrepancy, however: the axons of Mariani's 'blue cone' bipolars end in the proximal part of the inner plexiform layer, whereas the dendrites of De Monasterio's short-wave ganglion cells make contacts in the distal part of the inner plexiform layer.] In the lateral geniculate nucleus, the cells that draw excitatory inputs from the short-wave cones appear to be found predominantly in the parvocellular layers 3 and 4 (Schiller & Malpeli, 1978; Michael, 1988). The primordial subsystem may remain distinct as far as area 17 of the primate cortex: Ts'o & Gilbert (1988) have examined the properties of individual cells in the so-called 'cytochrome oxidase blobs' of layers 2 and 3 (Livingstone & Hubel, 1984; see Fig. 4) and report that cells within a given blob exhibited the same type of colour opponency, 'blue-yellow' or 'red-green', the latter type being three times as common as the former. It is likely that Ts'o & Gilbert's 'blue-yellow' cells correspond to what I am here calling the primordial subsystem; but it must be said that the cells were classified only by their responses to coloured lights, and the presence of short-wave inputs needs to be confirmed by the use of selective adaptation (e.g. Gouras, 1974) or of tritanopic substitution (Derrington *et al.* 1984). Since the maximum ratio of middle- to long-wave cone sensitivity occurs near 460 nm (see, for example, Boynton & Kambe, 1980), it is possible for the spectral cross-over point of a cell (the wavelength where an excitatory response replaces inhibition) to lie at short wavelengths even though the cell draws its opposed inputs from the middle- and long-wave cones.

#### *The primordial subsystem: conclusions*

Man and the Old World primates thus share with many mammals an ancient colour-vision channel, which depends on comparing the quantum catch of the short-wave cones with the quantum catch of a much more numerous class of cones, which have their peak sensitivity in the middle of the visible spectrum and which subserve the other functions of photopic vision. There may be a distinct morphological basis for this channel in the primate visual pathway, although the short-wave cones, and the cells that carry their signals, are always in the minority. It would be foolish to say that this primordial subsystem is a purely chromatic channel, since the several uses of colour vision (see above) themselves *require* that the chromatic signals carry with them a local sign; and the signals of this subsystem can sustain a number of spatial discriminations, provided only that the stimuli are of low spatial frequency. But the primordial colour channel has little role in the detection of edges and spatial detail.

#### *The second subsystem of colour vision*

As far as is known at present, the Old World primates are the only mammals

that share our own form of colour vision (Jacobs, 1981; Jacobs & Neitz, 1986), although the New World monkeys have found their private route to trichromacy (see below). The recent ancestors of the Old World primates acquired a second subsystem of colour vision, overlaid on the first; and the two subsystems remain relatively independent: physiologically independent at early stages of the visual system and psychophysically independent in detection and discrimination (Mollon & Jordan, 1989). The second subsystem depends on a comparison of the rates at which quanta are caught by the long- and middle-wave cones (Fig. 3); and this subsystem appears to have arisen through the duplication of a gene that coded for the photopigment of an ancestral middle-wave cone. (By 'middle-wave' is meant here only that the peak sensitivity lay in the range 510–570 nm, a range that includes the present 'long-wave' cone.) The recent differentiation of the long- and middle-wave photopigments has long been suspected (see e.g. Ladd-Franklin, 1892) on account of the distribution of trichromacy among the mammals (Jacobs, 1981) and the fact that hereditary disorders of colour vision chiefly affect the long- and middle-wave receptors (Pokorny *et al.* 1979). But convincing evidence has recently come from the molecular genetic results of Nathans and his collaborators (Nathans *et al.* 1986b), who have shown that the inferred amino-acid sequences of the middle- and long-wave human pigments are 96 % identical. Moreover, the two genes remain juxtaposed in a tandem array on the q-arm of the X-chromosome (Vollrath *et al.* 1988). The juxtaposition and the extreme homology of these genes render them vulnerable to misalignment when the X-chromosomes come together at meiosis; and Nathans and his collaborators suppose that the high incidence of human colour deficiency arises from the unequal crossing-over that can follow such misalignments (Nathans *et al.* 1986a).

#### *Co-evolution of fruit and of primate trichromacy?*

Interspecific variations occur in the photopigments of salmonid fish that have been isolated in land-locked glacial lakes for only a short evolutionary period (Bridges & Yoshikami, 1970; Bridges & Delisle, 1974), and it has become customary to suppose that opsins (the protein moieties of the photopigments) can evolve very rapidly. Moreover, Nathans *et al.* (1986a) have suggested how, in man, hybrid genes could readily be formed that coded for pigments with different spectral sensitivities (see above). The very consistency of the photopigments of catarrhine monkeys thus becomes a matter for remark. The monkeys so far examined – macaques (Bowmaker *et al.* 1978; Hárosi, 1982; Schnapf *et al.* 1987), baboons (Bowmaker *et al.* 1983), patas monkeys, talapoin, guenons (J. K. Bowmaker, S. Astell & J. D. Mollon, unpublished observations) – vary widely in their habitat, their size and their bodily colourings; yet they all exhibit a middle-wave pigment with a peak sensitivity close to 535 nm and a long-wave pigment with a peak sensitivity close to 565 nm. What this set of monkeys do have in common is that a substantial part of their diet consists of fruit; the proportion varies for different species, but is as high as 85 % in the case of the moustached guenon *Cercopithecus cephus* (Sourd & Gautier-Hion, 1986; Gautier-Hion, 1988).

Polyak (1957) explicitly proposed that the trichromatic colour sense of primates co-evolved with coloured fruits, such as mangoes, bananas, papaya, and the fruits of the citrus family. And recent ecological studies lend fresh credence to this view. Gautier-Hion *et al.* (1985) have examined the relationships between frugivores and fruits in a tropical rain forest, and it becomes clear that it is not enough to describe an animal as fruit-eating. There exists a category of fruit that is disproportionately taken by monkeys: such fruits are orange or yellow in colour, weigh between 5 and 50 g, and are either dehiscent with arillate seeds or are succulent and fleshy. In contrast, the fruits predominantly taken by birds are red or purple, and are smaller, while the fruits taken by ruminants, squirrels and rodents are dull-coloured (green or brown) and have a dry fibrous flesh.

Thus, there exist fruits that appear to be specialized for attracting monkeys. The monkey serves the fruiting tree by dispersing the seed in one of two ways. When the seed is large and the soft flesh is free from the seed, the latter is often spat out; this usually happens at some distance from the parent tree, because the monkeys fill their cheek-pouches and move to another place to eat the contents. When the seeds are small, they are usually swallowed and are excreted intact. Thus, the monkey is a disperser for the tree, rather than a predator (Gautier-Hion *et al.* 1985). To find orange and yellow fruits amongst foliage, the monkey needs trichromatic colour vision; without it, the monkey would be at the very disadvantage emphasized in the early accounts of human colour blindness (see above).

It would be instructive to know whether the leaf-eating catarrhine monkeys differ from the frugivores, but nothing is yet known about their colour vision.

#### *The second subsystem and the parvocellular pathway*

At the early stages of the visual system, there does not appear to be a channel devoted exclusively to carrying the second type of chromatic information, the information obtained by comparing the quantum catches of long- and middle-wave cones; rather the second subsystem is parasitic upon an existing channel that carries information about spatial detail. The substrate of this channel is the  $P\beta$  cell of the primate retina (Perry & Cowey, 1981) and the predominant type of unit in the parvocellular layers of the lateral geniculate nucleus (represented in red in Fig. 4). The receptive fields of such cells are divided into antagonistic centre and surround regions and, thus, the cells are very sensitive to spatial contrast. In the case of the Old World primates, the centre input to such cells is drawn from one class of cone (often, perhaps, from one individual cone), whereas the surround is drawn either from a different class of cone, or promiscuously from middle- and long-wave cones (Lennie *et al.* 1989). Thus the response of any individual cell is ambiguous: as Ingling & Martinez (1983) have emphasised, the cell will be colour-specific at low spatial frequencies but will respond to all wavelengths at higher spatial frequencies.

Shapley & Perry (1986) have argued that the primary function of the parvocellular system is colour vision, and that the analysis of spatial contrast depends



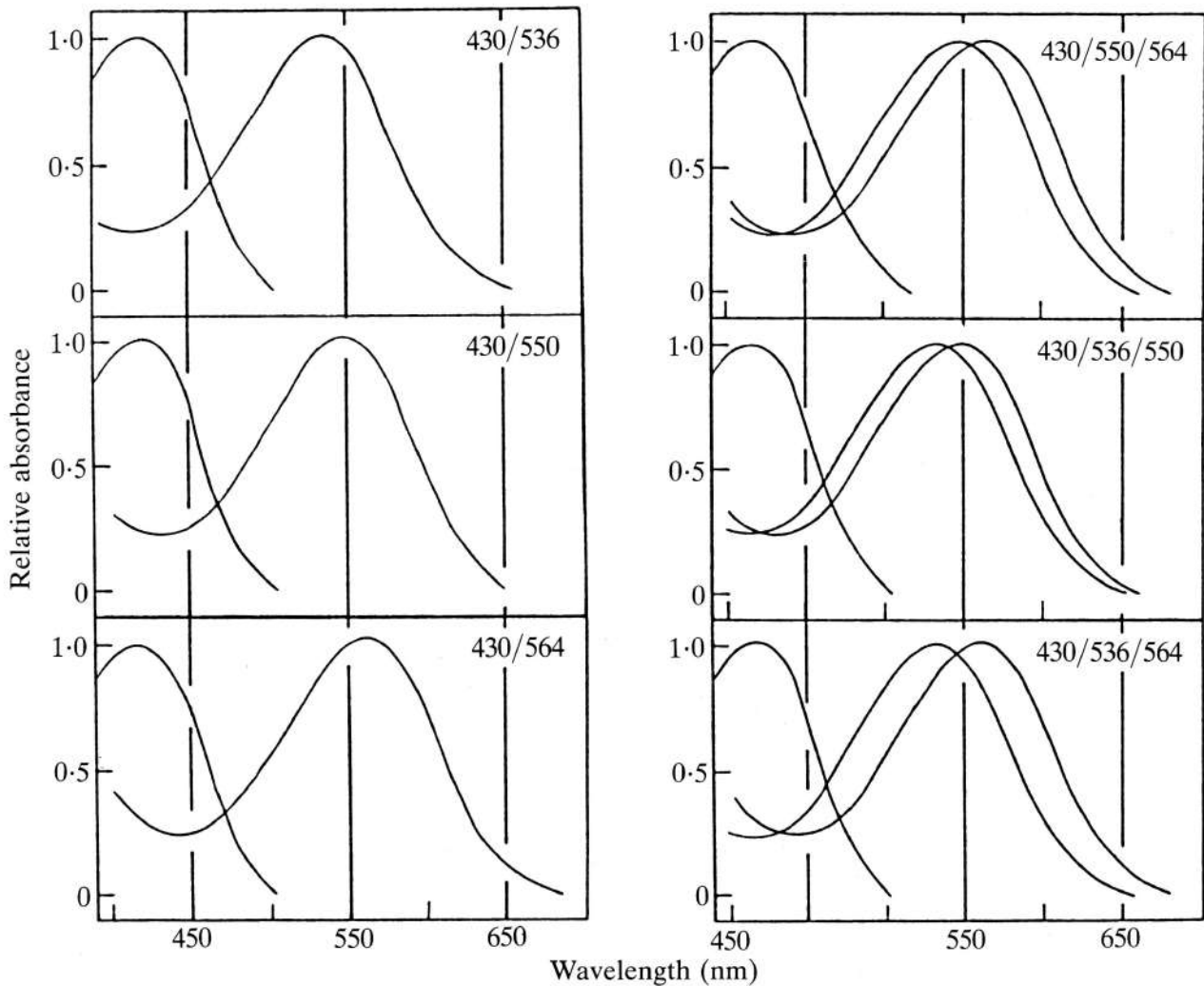


Fig. 5. The photopigments of six types of squirrel monkey (*Saimiri sciureus*). Each panel shows the absorbance spectra of the photopigments found microspectrophotometrically in a given phenotype. To the left are shown the three possible combinations of pigments that occur in dichromat phenotypes, that is, all males and some females. To the right are shown the combinations that occur in the three possible trichromatic phenotypes, which are always female.

predominantly on the magnocellular pathway. Arguments against this position are given by Mollon & Jordan (1989).

### The polymorphism of colour vision in platyrrhine primates

Until recently the New World monkeys were held to be protanopes, and to represent an earlier stage of our own colour vision. We now know that platyrrhine colour vision is *sui generis* and is characterized by striking within-species variability (Jacobs, 1983). In a double-blind study that combined behavioural testing and microspectrophotometry (Mollon *et al.* 1984; Bowmaker *et al.* 1987), six phenotypes were identified in the squirrel monkey (*Saimiri sciureus*): all males and some females were dichromats, combining a short-wave pigment with one of three possible pigments in the red–green spectral region (Fig. 5), whereas other females were trichromatic, combining a short-wave pigment with two pigments in the red–green region. These findings have been explained by a genetic model (Mollon



*et al.* 1984) that makes the following assumptions. (i) The squirrel monkey has only a single genetic locus for a pigment in the red–green spectral region. (ii) There are at least three alleles that can occur at this locus, the three alleles corresponding to three slightly different versions of the opsin of the photopigment. (iii) The locus is on the X-chromosome. (iv) In those females that are heterozygous at this locus, only one of the two alleles is expressed in any given cone cell, owing to the phenomenon of Lyonization or X-chromosome inactivation.

By this account, at least three photopigments are potentially available to the squirrel monkey in the red–green region. Male monkeys, it is supposed, can draw only one pigment from the set, since they have only one X-chromosome; so males are obligatory dichromats. Females may draw either one or two pigments from the set; if they inherit the same allele from both parents, they will be dichromatic, but if they inherit two different alleles, they will be trichromatic. We must make the explicit assumption that the visual system of the heterozygous female is sufficiently plastic to exploit the presence of distinct subsets of signals from the two types of cone with peak sensitivities in the red–green region. And it may be the heterozygous female that leads her troop in the daily search for ripe fruit.

I am grateful to Dr M. Webster and Mr J. Reffin for comments on the text and to Gabriele Jordan for the design of Fig. 4.

### References

- BARLOW, H. B. (1981). Critical limiting factors in the design of the eye and visual cortex. *Proc. R. Soc. B* **212**, 1–34.
- BARLOW, H. B. (1982). What causes trichromacy? A theoretical analysis using comb-filtered spectra. *Vision Res.* **22**, 635–643.
- BOWMAKER, J. K., DARTNALL, H. J. A., LYTHGOE, J. N. & MOLLON, J. D. (1978). The visual pigments of rods and cones in the rhesus monkey, *Macaca mulatta*. *J. Physiol., Lond.* **272**, 329–348.
- BOWMAKER, J. K., JACOBS, G. H. & MOLLON, J. D. (1987). Polymorphism of photopigments in the squirrel monkey: a sixth phenotype. *Proc. R. Soc. B* **231**, 383–390.
- BOWMAKER, J. K., MOLLON, J. D. & JACOBS, G. H. (1983). Microspectrophotometric results for Old and New World primates. In *Colour Vision: Physiology and Psychophysics* (ed. J. D. Mollon & L. T. Sharpe), pp. 57–68. London: Academic Press.
- BOYLE, R. (1688). *Some Uncommon Observations About Vitiated Sight*. London: J. Taylor.
- BOYNTON, R. M. (1979). *Human Color Vision*. New York: Holt, Rinehart & Winston.
- BOYNTON, R. M. & KAMBE, N. (1980). Chromatic difference steps of moderate size measured along theoretically critical axes. *Color Res. Applic.* **5**, 12–23.
- BRIDGES, C. D. B. & DELISLE, C. E. (1974). Evolution of visual pigments. *Expl Eye Res.* **18**, 323–332.
- BRIDGES, C. D. B. & YOSHIKAMI, S. (1970). Distribution and evolution of visual pigment in salmonid fishes. *Vision Res.* **10**, 609–626.
- BRINDLEY, G. S. (1954). The summation areas of human colour-receptive mechanisms at increment threshold. *J. Physiol., Lond.* **124**, 400–408.
- CHISHOLM, J. J. (1869). Colour blindness, an effect of neuritis. *Ophthalmic Hosp. Rep.*, April 214–215.
- COLQUHOUN, H. (1829). An Account of Two Cases of Insensibility of the Eye to certain of the Rays of Colour. *Glasgow Medical Journal* **2**, 12–21.
- COOMBE, G. (1824). Case of Deficiency in the Power of Perceiving and Distinguishing Colours,

- accompanied with a small development of the Organ, in Mr. James Milne, Brassfounder in Edinburgh. *Transactions of the Phrenological Society*, pp. 222–234.
- DAMASIO, A., YAMADA, T., DAMASIO, H., CORBETT, J. & MCKEE, J. (1980). Central achromatopsia: Behavioral, anatomic and physiologic aspects. *Neurology* **30**, 1064–1071.
- DARTNALL, H. J. A., BOWMAKER, J. K. & MOLLON, J. D. (1983). Human visual pigments: microspectrophotometric results from the eyes of seven persons. *Proc. R. Soc. B* **220**, 115–130.
- DE MONASTERIO, F. M. (1979). Asymmetry of on- and off-pathways of blue-sensitive cones of the retina of macaques. *Brain Res.* **166**, 39–84.
- DERRINGTON, A. M., KRAUSKOPF, J. & LENNIE, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol., Lond.* **357**, 241–265.
- GAUTIER-HION, A. (1988). The diet and dietary habits of forest guenons. In *A Primate Radiation: Evolutionary Biology of the African Guenons* (ed. A. Gautier-Hion, F. Bourlière, J.-P. Gautier & J. Kingdon), pp. 257–283. Cambridge: Cambridge University Press.
- GAUTIER-HION, A., DUPLANTIER, J.-M., QURIS, R., FEER, F., SOURD, C., DECOUX, J.-P., DUBOST, G., EMMONS, L., ERARD, C., HECKETSWEILER, P., MOUNGAZI, A., ROUSSILHON, C. & THIOLLAY, J.-M. (1985). Fruit characteristics as a basis of fruit choice and seed dispersal in a tropical forest vertebrate community. *Oecologia* **65**, 324–337.
- GOURAS, P. (1974). Opponent-colour cells in different layers of foveal striate cortex. *J. Physiol., Lond.* **238**, 582–602.
- GOURAS, P. & EGGERS, H. (1983). Responses of primate retinal ganglion cells to moving spectral contrast. *Vision Res.* **23**, 1175–1182.
- HÁROSI, F. I. (1982). Recent results from single-cell microspectrophotometry: cone pigments in frog, fish and monkey. *Color Res. Appl.* **7**, 135–141.
- HUDDART, J. (1777). An account of Persons who could not distinguish Colours. *Phil. Trans. R. Soc.* **67**, 260–265.
- INGLING, C. R. & MARTINEZ, E. (1983). The spatiochromatic signal of the r-g channel. In *Colour Vision: Physiology and Psychophysics* (ed. J. D. Mollon, & L. T. Sharpe), pp. 433–444. London: Academic Press.
- JACOBS, G. H. (1981). *Comparative Color Vision*. New York: Academic Press.
- JACOBS, G. H. (1983). Within-species variations in visual capacity among squirrel monkeys (*Saimiri sciureus*). *Vision Res.* **23**, 1267–1277.
- JACOBS, G. H. & NEITZ, J. (1986). Spectral mechanisms and color vision in the tree shrew (*Tupaia belangeri*). *Vision Res.* **26**, 291–298.
- KALMUS, H. (1955). The familial distribution of congenital tritanopia, with some remarks on some similar conditions. *Annals hum. Genet.* **20**, 39–56.
- KATZ, D. (1935). *The World of Colour*. London: Kegan Paul, Trench, Trubner & Co.
- KINGDON, J. S. (1980). The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *Trans. zool. Soc. Lond.* **35**, 425–475.
- KINGDON, J. S. (1988). What are face patterns and do they contribute to reproductive isolation in guenons? In *A Primate Radiation: Evolutionary Biology of the African Guenons* (ed. A. Gautier-Hion, F. Bourlière, J.-P. Gautier, & J. Kingdon), pp. 227–245. Cambridge: Cambridge University Press.
- KOFFKA, K. (1935). *Principles of Gestalt Psychology*. London: Kegan Paul, Trench, Trubner & Co.
- LADD-FRANKLIN, C. (1892). A new theory of light sensation. In *International Congress of Psychology, 2nd Congress, London, 1892*. Kraus Reprint, 1974.
- LAND, E. (1983). Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5163–5169.
- LENNIE, P., HAAKE, P. W. & WILLIAMS, D. R. (1989). Chromatic opponency through random connections to cones. *Investig. Ophthalmol. vis. Sci.*, **30** (Suppl.), 322.
- LIEBMANN, S. (1927). Ueber das Verhalten farbiger Formen bei Helligkeitsgleichheit von Figur und Grund. *Psychol. Forschung.* **9**, 300–353.
- LIVINGSTONE, M. & HUBEL, D. (1984). Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.* **4**, 309–356.

- LIVINGSTONE, M. & HUBEL, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* **240**, 740–749.
- MCDUGALL, W. (1901) Some new observations in support of Thomas Young's theory of light and colour-vision (II). *Mind* **10**, 210–245.
- MARIANI, A. P. (1984). Bipolar cells in monkey retina selective for the cones likely to be blue-sensitive. *Nature, Lond.* **308**, 184–186.
- MEADOWS, J. C. (1974). Disturbed perception of colours associated with localized cerebral lesions. *Brain* **97**, 615–632.
- MICHAEL, C. R. (1988). Retinal afferent arborization patterns, dendritic field orientations, and the segregation of function in the lateral geniculate nucleus of the monkey. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4914–4918.
- MOLLON, J. D. (1980). Post-receptor processes in colour vision. *Nature, Lond.* **283**, 623–624.
- MOLLON, J. D. (1982). A taxonomy of tritanopias. In *Colour Vision Deficiencies*, VI (ed. G. Verriest) (*Documenta Ophthalmologica Proceedings Series* **33**), pp. 87–101. The Hague: Dr W. Junk.
- MOLLON, J. D. (1989). On the nature of models of colour vision. *Die Farbe* (in press).
- MOLLON, J. D., BOWMAKER, J. K. & JACOBS, G. H. (1984). Variations of colour vision in a New World primate can be explained by polymorphism of retinal photopigments. *Proc. R. Soc. B* **222**, 373–399.
- MOLLON, J. D., ESTEVEZ, O. & CAVONIUS, C. R. (1989). The two subsystems of colour vision and their rôles in wavelength discrimination. In *Vision: Coding and Efficiency* (ed. C. B. Blakemore). Cambridge: Cambridge University Press (in press).
- MOLLON, J. D. & JORDAN, G. (1989). Eine evolutionaere Interpretation menschlichen Farbsehens. *Die Farbe* (in press).
- MOLLON, J. D., NEWCOMBE, F., POLDEN, P. G. & RATCLIFF, G. (1980). On the presence of three cone mechanisms in a case of total achromatopsia. In *Colour Vision Deficiencies*, V (ed. G. Verriest), pp. 130–135. Bristol: Adam Hilger.
- MORGAN, M. J., MOLLON, J. D. & ADAM, A. (1989). Dichromats break colour-camouflage of textural boundaries. *Investig. Ophthalmol. vis. Sci.* **30** (Suppl.), 220.
- NATHANS, J., PIANTANIDA, T. P., EDDY, R. L., SHOWS, T. B. & HOGNESS, D. S. (1986a). Molecular genetics of inherited variation in human color vision. *Science* **232**, 203–210.
- NATHANS, J., THOMAS, D. & HOGNESS, D. S. (1986b). Molecular genetics of human color vision: the genes encoding blue, green and red pigments. *Science* **232**, 193–202.
- NICHOLL, W. (1818). Account of a case of defective power to distinguish colours. *Medico-Chirurgical Trans.* **9**, 359–363.
- PERRY, V. H. & COWEY, A. (1981). The morphological correlates of X- and Y-like retinal ganglion cells in the retina of monkeys. *Expl Brain Res.* **43**, 226–228.
- POKORNY, J., SMITH, V. C., VERRIEST, G. & PINCKERS, A. J. L. G. (eds) (1979). *Congenital and Acquired Color Vision Defects*. New York: Grune & Stratton.
- POLYAK, S. (1957). *The Vertebrate Visual System*. Chicago: Chicago University Press.
- RODIECK, R. W. (1988). The primate retina. In *Comparative Primate Biology* (ed. H. D. Stecklis & J. Erwin), pp. 203–278. New York: Alan R. Liss Inc.
- RUSHTON, W. A. H. (1972). Pigments and cones in colour vision. *J. Physiol., Lond.* **220**, 1P–31P.
- SCHILLER, P. H. & MALPELLI, J. G. (1978). Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *J. Neurophysiol.* **41**, 788–797.
- SCHNAPF, J. L., KRAFT, T. W. & BAYLOR, D. A. (1987). Spectral sensitivity of human cone receptors. *Nature, Lond.* **325**, 439–441.
- SHAPLEY, R. & PERRY, V. H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends Neurosci.* **9**, 229–235.
- SOURD, C. & GAUTIER-HION, A. (1986). Fruit selection by a forest guenon. *J. Animal Ecol.* **55**, 235–244.
- STILES, W. S. (1949). Increment thresholds and the mechanisms of colour vision. *Documenta Ophthalmologica* **3**, 138–163.
- STILES, W. S. (1952). Colour vision: a retrospect. *Endeavour* **11**, 33–40.
- STILLING, J. (1883). *Pseudo-isochromatische Tafeln fuer die Pruefung des Farbennes*. Kassel, Berlin: Theodor Fischer.

- STROMEYER, C. S., KRONAUER, R. E., MADSEN, J. C. & COHEN, M. A. (1980). Spatial adaptation of short-wavelength pathways in humans. *Science* **207**, 555–557.
- STROMEYER, C. S. & LEE, J. (1987). Motion and short-wave cone signals: influx to luminance mechanisms. *Investig. Ophthalmol. vis. Sci.* **28** (Suppl.), 232.
- TANSLEY, B. W. & BOYNTON, R. M. (1976). A line, not a space, represents visual distinctness of borders formed by different colors. *Science* **191**, 954–957.
- THOMA, W. & SCHEIBNER, H. (1980). Die spektrale tritanopische Saettigungsfunktion beschreibt die spektrale Distinktibilitaet. *Farbe Design* **17**, 49–52.
- TS'O, Y. & GILBERT, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *J. Neurosci.* **8**, 1712–1727.
- VOLLRATH, D., NATHANS, J. & DAVIES, R. W. (1988). Tandem array of human visual pigment genes at Xq28. *Science* **240**, 1669–1672.
- WERTHEIMER, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Pyschol. Forschung.* **4**, 301–350.
- ZEKI, S. M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *J. Physiol., Lond.* **277**, 273–290.