J. Physiol. (1972), **220**, *pp.* 1–31*P*
With 16 *text-figures*
Printed in Great Britain

REVIEW LECTURE

PIGMENTS AND SIGNALS IN COLOUR VISION

INVITED LECTURE TO THE PHYSIOLOGICAL SOCIETY

By W. A. H. RUSHTON
From Trinity College, Cambridge

Given at the meeting of the Physiological Society at Chelsea College, London on 17 April 1970

The trouble with colour vision is the mentality of those that write on it – and of those that read. In most aspects of physiology it is sufficient to offer a fairly plausible and adequate hypothesis; but colour visionaries want nothing less than the truth. The cause of this unreasonable demand lies in this, that whereas nearly all the phenomena of nature are simply observed, those of sensory physiology can also be experienced. So in colour vision we perceive the essential hollowness of formal scientific explanation. The considerations of this lecture, I am afraid, will remain hollow, for the filling is such stuff as dreams are made on, and is not very suitable for the Journal of Physiology.

The eye has long been recognized as a camera, first as a *camera obscura*, then as a photographic camera, and now as a television camera. Obviously the latter is a better model, in that the picture is not pickled on the retina, but transmitted in code to a distant part for immediate computing. The physics of transmission in a million nerve fibres, of course, is entirely different from radio waves modulated within a narrow band-width, but both systems transmit information from which much of the original optical image could be reconstructed in form and colour.

The first stage in this long and highly complex process is the transduction of light; the impalpable optical image on the retina is turned into a more substantial picture of chemical change. The next stage is the way that one or two molecules of chemical change can cause nerve excitation. Following this is endless processing of nerve signals at various levels of the visual system. In Part I of this Lecture I shall speak mainly of quantum catching – what light is caught by various photoreceptors; that century-old question now at last is settled. In Part II I shall touch speculatively upon the nerve signals – the growing point in vision; ten years from now our guesses will be replaced by knowledge.

Part I. The catch of quanta

Young's explanation of trichromacy

In 1801 Thomas Young, aged 28, physician of St George's Hospital, gave the Bakerian Lecture to the Royal Society. He spoke on the nature of light, gave his famous proof that it was a wave motion, measured the wave-length by his interference fringes and threw out a suggestion as to the nature of colour vision – a master-suggestion, for it resolved a paradox that had puzzled everyone for a century (Young, 1802).

Newton in 1666 at the age of 23 had performed his famous prism experiment and shown that sunlight consisted of an infinite number of rays each bent differently by a prism and hence falling in a different place on the opposite wall of his room in Trinity College. Each of these rays appeared to have a different colour, and each was elementary (or monochromatic) in the sense that each, when passed again through a prism, was not further changed either in colour or in refrangibility (Newton, 1672). It thus was clear that the retinal image of any natural scene is the superposition of multiple images formed by the infinite variety of coloured rays. It seemed to follow that painters, in order to get a life-like representation, would need an infinite variety of coloured paints. But the great painters, say, of the Italian Renaissance, achieved their master representations not by using an enormous variety of paints but by delicate mixtures upon the palette of quite a few bright paints – indeed only three paints seemed theoretically necessary. During the eighteenth century it gradually became accepted that colour vision was trichromatic in the sense that any colour could be produced by a mixture of three primary colours, yet Lomonosov's (1757) view that there were just three kinds of light could not be correct, for Newton had shown that there was an infinite number. This was the paradox that was resolved by Young's pregnant aside in the course of his Lecture.

Young was fascinated by the way that the eye collects and analyses visual information. He was elected a Fellow of the Royal Society at the age of 21 for showing how the eye accommodated for objects at various distances, and he knew that light from every point in the outside world was brought to a focus at a distinct point on the retina. It was unreasonable to suppose that each of these retinal points could do for light what it takes the whole cochlea to do for sound, namely to perform a frequency analysis and encode the whole spectrum in a range of nerve messages. Nor indeed can the eye do this; it needs Newton's prism (or Noah's rainbow). Young wrote, 'Now, as it is almost impossible to conceive each sensitive point of the retina to contain an infinite number of particles each vibrating in perfect unison with every possible undulation, it becomes necessary to suppose

the number limited; for instance to the three principal colours red, yellow and blue...'.

Both Young and Newton saw clearly that when we mix (say) red and yellow to make an orange which to the eye is indistinguishable from spectral orange, the physical lights red + yellow have not been changed into spectral orange. Looked at through a prism they are still seen to be what they always were, 'red + yellow' on the one hand 'spectral orange' on the other. They look the same because we are deceived. For Newton, the physicist, the nature of light was fundamental; the way that the eye can be deceived, not so important. But Young was also a physician, and he realized that the eye's limitations provide the key to its mode of action. That colour vision was limited to trichromacy suggested to him a mechanism, which as far as it went, is exactly right.

Young, in his famous suggestion, pointed out that the trichromacy of colour vision could be reconciled with the infinite variety of light species if there are three resonators in the eye, each 'capable of being put into motion less or more forcibly by undulations differing less or more from a perfect unison...'. He did not know what those resonators were, but he knew the velocity of light from Roemer and from Bradley and he knew its wave-length from his own interference fringes; thus he knew that the resonators vibrated at about 10^{14} cycles per second. This is much too fast for any material object; it must be something going on in the atoms of his famous colour-blind contemporary John Dalton. In fact the resonators are electrons in the π orbitals of cone pigment molecules. Enough of that.

The visual pigments

If light is to have any material effect it must transfer its energy to matter. This occurs in the retina when light is absorbed by a photosensitive pigment, a chemical like that on a photographic film, in which the energy absorbed is not simply degraded into heat (as in the pigments of pictures) but is converted into chemical change. Boll (1876) noticed that light falling on frog's retina (hitherto screened) caused the pink colour to fade, and Kühne (1878) discovered a great deal of what we now know of the bleaching and regeneration of this pink pigment, *rhodopsin*, which can be seen lying in the outer segments of the rods.

A pigment which reacts to light and is situated in the rods is very possibly a *visual* pigment, the photopigment of rod vision. If so, it must satisfy a rather strict test: lights of different wave-lengths, adjusted in energy so that they are equal for rod vision, should be found to be equally absorbed by the pigment. This was first established by Koenig (1894) and has been confirmed with increasing sophistication many times since Crescitelli & Dartnall (1953). So rhodopsin is a visual pigment, and rods

are excited by the quanta it catches. Rod vision, however, is twilight vision which is colourless. Thus, though the rod input has two variables, wave-length and energy, the output differs only in one respect, namely 'brightness'. The effect of wave-length (as Koenig showed) is simply to modify the proportion of light that is absorbed, i.e. the proportion of incident quanta that are caught. Every quantum that *is* caught, however (whatever its wave-length), produces the same effect – its unit contribution to output. This very important property of rods, and indeed also of each kind of cone, this limitation of output to a single dimension of change, may be called the *Principle of Univariance* and stated thus: 'The output of a receptor depends upon its quantum catch, but not upon what quanta are caught.

Since rod vision is colourless, cones must be involved in colour vision. It has often been argued that rods also play a part in normal colour vision. The view that they are the blue receptors has an interesting range of phenomena to support it (Willmer, 1946, 1961). But it cannot be right, since rods are equally sensitive to rays of light entering the eye through any point of the pupil, whereas cones are much more sensitive to central rays (Stiles–Crawford effect). Blue receptors are more sensitive to the central rays and hence they cannot be rods. There is a rare class of colour defective whose retinal receptors consist only of rods and blue cones. In a recent study of these cases, Alpern, Lee, Maaseidvaag & Miller (1971) contrast expertly the properties of the rods and the blue cones; this should finally bury the conjecture that they are identical. Rods, however, certainly can contribute to colour vision at rather low levels (in parafoveal vision); their effect is to desaturate, i.e. to dilute cone colour by adding white.

Young's theory of colour vision may now be stated in terms of cone pigments. 'There are three classes of cone each containing a different visual pigment. The output of each cone is univariant, depending simply upon the quantum catch of its pigment. Our sensation of colour depends upon the ratios of these three cone outputs'. All of these statements are known to be true.

The cone pigments

The first definitive measurements of visual pigment in the outer segments of single cones was achieved in the goldfish by Marks (1965). Three classes of cone were found each with a different pigment, which could not be expected to be the same as those in man since the pigments of fresh water fish are compounds of vitamin A_2 whereas human pigments are compounded with A_1. Marks' highly sophisticated technique, however, was applied to isolated retinas or cones from man and monkeys (Marks,

Dobelle & MacNichol, 1964) and the results obtained are shown in Fig. 1. The authors left open the possibility that the pigments measured in single cones might be mixtures, but this is perhaps unlikely in view of the good coincidence of their points with the black and white circles in Fig. 1. These are plotted from objective pigment measurements made by reflexion densitometry on the living eyes of normals and of colour defectives (Baker & Rushton, 1965, Fig. 6). One type of defective, the protanope, has only

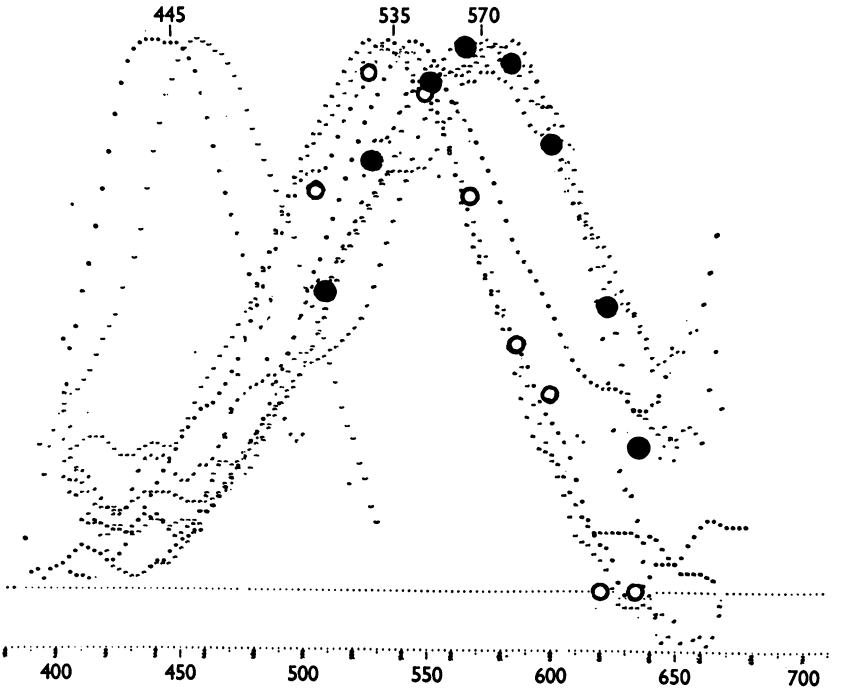

Fig. 1. Absorption spectrum of single cones in excised retinas from man and monkey. Abscissae, wave-length (nm), ordinates absorption plotted by computer from photocell output, programmed to compensate for various factors and plot scaled to a fixed maximum. Each point is a separate measurement, each dotted curve is from a separate cone (from Marks, Dobelle & MacNichol, 1964, copyright 1964 by the American Association for the Advancement of Science.) Heavy black and white circles represent similar measurements by reflexion densitometry in the living eye of normal man (from Baker & Rushton, 1965.)

the green-catching pigment *chlorolabe* sensitive in the red-green spectral range (Rushton, 1963), and the fact that its spectral sensitivity (white circles) coincides with that of single 'green' cones is evidence that these contain chlorolabe and that only. Similarly the other kind of common colour defective, the deuteranope, has only the pigment *erythrolabe* (= red

catching) sensitive in the red-green range (Rushton, 1965*a*); the fact that its spectral sensitivity (black circles) coincides with that of single 'red' cones is evidence that these contain erythrolabe and that only. It thus appears that each cone contains only one pigment and that 'red' cones are lacking or without response in the protanope, 'green' cones lacking in the deuteranope and by analogy, that 'blue' cones are lacking in the tritanope.

Colour defectives

The statement just made about the pigment deficiencies in red-green defectives is so important that it needs supporting evidence.

About 8% of males have some genetic defect in their red-green colour mechanism appearing as a sex-lined recessive character. In about a quarter of these, the defect is so pronounced that the whole spectrum can be matched with two lights only, say a red and blue. They are therefore called 'dichromats'; the remainder are 'anomalous trichromats'. A further fact sharply divides the dichromats; although they can all match a red light exactly with a yellow, the intensity of yellow that looks the same as a fixed red is very different. *Deuteranopes* require the yellow to be about as bright as the red (as judged by normal eyes) whereas *protanopes* who are red-blind, match a good red by a very dim and dirty yellow.

Now the amount of pigment in the foveal cones of any living person can be measured by the technique of retinal densitometry. Light is shone on to the fovea in an ophthalmoscopic apparatus, and the fraction reflected back from the *fundus* is deflected on to a photocell and measured. This light has passed through the retina and back and must have suffered absorption by the cone pigments there. Consequently any change in pigment density (e.g. bleaching or regeneration) will result in a change in the light emerging from the eye, and can be measured by the photocell upon which it falls (Rushton, 1956; Hood & Rushton, 1971). Fig. 2 shows cone pigment measurements made in this way upon a protanope (triangles) and a deuteranope (circles), using light of various wave-lengths to examine the change in the spectrum, following up the earlier work of Rushton (1955).

Black triangles show the change in density of the protanope's cones following 50% bleaching by a bright red light. White triangles show the same thing (after a regeneration period in the dark) when the bleaching light was blue–green and of such strength that it bleached the same amount, measured at 555 nm. It is seen that black and white triangles coincide not only at 555 but at every wave-length. This does not happen with the normal eye and could not happen in any eye that possessed both a red- and a green-sensitive pigment. For the red bleaching light would bleach the red-sensitive pigment more and produce more change in the red spectral region, and the blue–green bleaching light would produce

more change in the green. Consequently, if the protanope contained two pigments in the red–green range we should see black triangles lying above white on the right, and below them on the left. The fact that this is never found means that protanope's have only one pigment and consequently only one kind of resulting curve whatever the wave-length of the bleaching light (univariance). The same experiments were also performed upon the

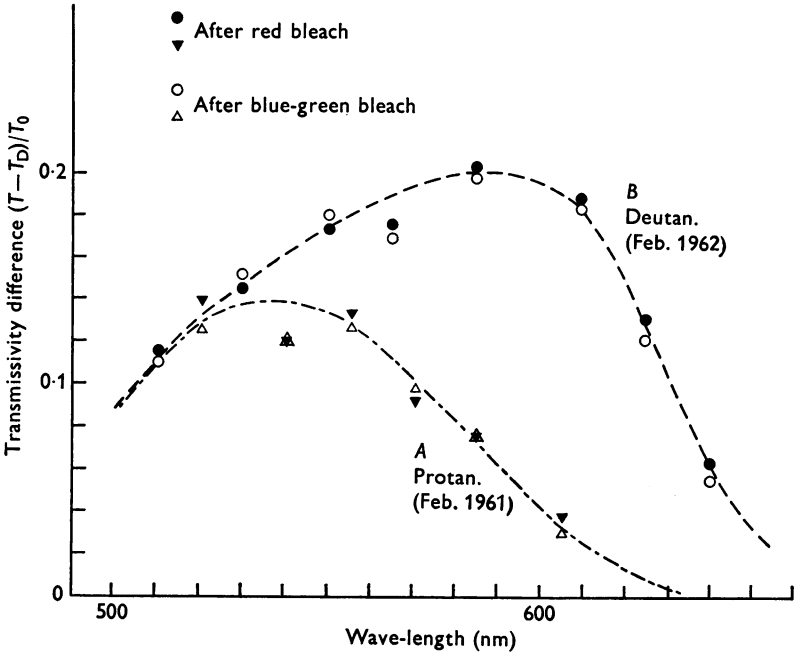

Fig. 2. Curve *A* plots for a protanope the increase in reflected light when the fovea was half-bleached by red light (black triangles) or by blue-green light (white triangles). Curve *B* plots the same for a deuteranope after half-bleaching with a red light (black circles) or a blue-green light (white circles). The coincidence of black and white symbols means that only one cone pigment is present; the separation of circles from triangles means that the deuteranope's pigment is more red-sensitive than the protanope's (from Rushton, 1965*b*).

deuteranope and the same coincidence found between white and black circles (Fig. 2). Thus deuteranopes also have only one cone pigment in the red–green range. But curve *B* is quite a different shape from *A* with far greater absorption in the red. So protanopes are 'red-blind' because they have no erythrolabe to catch 'red' quanta; but deuteranopes, though they lack chlorolabe are not really 'green-blind' since curve *B* absorbs in the green about as strongly as curve *A* does.

We have been assuming that chlorolabe and erythrolabe, the pigments

we measure, are *visual* pigments, i.e. those whose quantum catch leads to vision. If so, light of various wave-lengths, adjusted in intensity so that each bleaches the pigment at the same rate, will all be judged by the dichromat to be equally bright. Rushton (1963, 1965*a*) found this to be so, and the results of somewhat better experiments (Mitchell & Rushton,

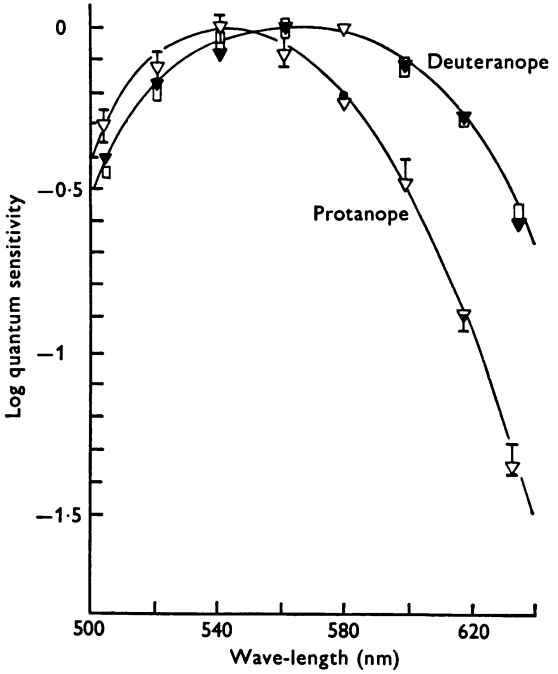

Fig. 3. The curves show the log spectral sensitivity of protanopes (left) and deuteranopes (right). Vertical bars or rectangles show minus the log quantum energies (± S.E.) of lights of various wave-lengths which all bleach chlorolabe or erythrolabe at the same rate. Triangles show minus the log energies (± S.E.) that this dichromat finds equally bright. Coincidence means that lights that look equally bright bleach at equal rates, thus the pigment measured is that responsible for brightness (from Mitchell & Rushton, 1972*a*).

1972*a*) are shown in Fig. 3. The triangles plot, for various wave-lengths, the light energy, which looks equally bright to the deuteranope (black triangles) or to the protanope (white triangles). The rectangles on the deuteranope curve show the energies which bleach erythrolabe at equal rates; the lines on the protanope curve show the same for chlorolabe. Coincidence means that the energies for bleaching and for brightness are equivalent and hence that the pigment we measure is the one whose quantum catch leads to vision.

Normal subjects in the red–green range have two cone pigments; dichromats have lost one of these entirely. Is the pigment they possess the other normal pigment? Fig. 4 (from Mitchell & Rushton, 1972*b*) shows by white triangles the log spectral sensitivity of erythrolabe in the deuteranope (inverted triangles) and of the red cones in normals (erect triangles). Likewise the black triangles show the same for chlorolabe in the protanope (inverted) and the green cones in normals (erect). Their coincidence means that erythrolabe and chlorolabe are normal pigments, and that dichromats have only suffered the simple loss of one pigment, without modification of the other.

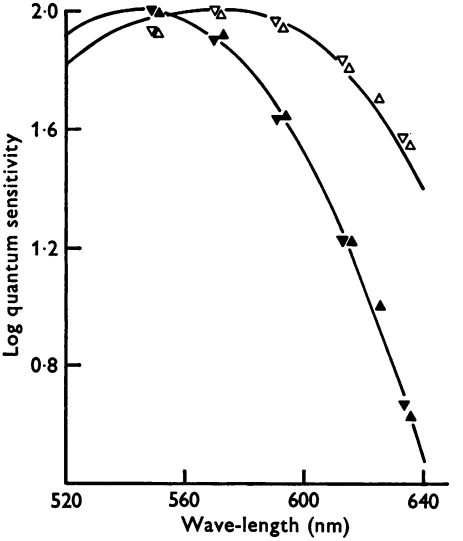

Fig. 4. The curves show the log spectral sensitivity of protanopes (left) and deuteranopes (right). Inverted triangles plot the minus log quantum energies (± 2 S.E.) of lights with various wave-lengths required to match a fixed yellow light. Erect triangles plot the same for the red pigment (white triangles) or the green pigment (black triangles) in the normal eye. Coincidence of erect and inverted triangles means that those pigment which the dichromats possess are normal pigments (from Mitchell & Rushton, 1972*b*).

Colour matching

The best books on colour vision devote a great deal of space and expertise to the question of colour matching. This is because we possess a huge body of exquisite colour matching results which in accuracy and stability transcend anything else in colour vision. Moreover there is a rigid mathematical technique for transforming all this material to meet any required situation. And historically it has seemed the best way to investigate the nature of Young's three resonators.

To physiologists who want only to understand roughly how lights can

be matched by mixing, it will probably be sufficient simply to consider the results of Fig. 1, or more clearly Fig. 5, where the same curves are represented, scaled not to have equal heights but equal areas between curve and base line. The ordinate drawn at any wave-length λ_1 will cut the curves R, G, B at heights r_1, g_1, b_1, and consequently the relative quantum catch

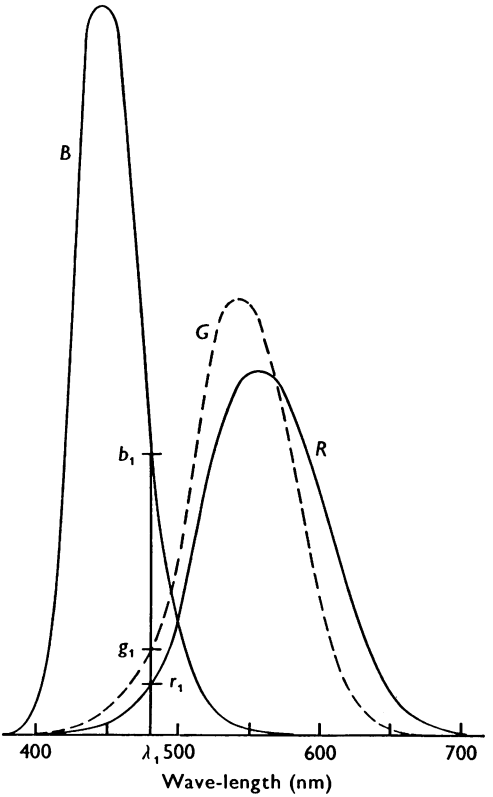

Fig. 5. The spectral sensitivity of red, green and blue cones (R, G, B). The curves are essentially a replot of those in Fig. 1 but scaled so that the area between the curve and the base line is the same for each.

in these cones is as $r_1 : g_1 : b_1$. If the light energy is made n_1 times as strong, obviously the catches become $n_1 r_1$, $n_1 g_1$, $n_1 b_1$. If now we add to this energy a second light of energy n_2 and wave-length whose ordinates on Fig. 5 are r_2, g_2, b_2, then the total quantum catch in R cones will be $(n_1 r_1 + n_2 r_2)$ with a similar expression for the green and blue cones.

Since any light whatever is made up of monochromatic components, the total quantum catch of the R cones may be calculated simply by adding the catches due to each component,

$$\text{thus total quantum catch} = n_1 r_1 + n_2 r_2 + n_3 r_3 + \ldots \qquad (1)$$

and the same may be done for the catches in G and B cones. If any two light mixtures have the same total quantum catch for R, and also the same for G and also the same for B, then the inputs will be identical in quantum catch for every cone, and hence by the principle of univariance the outputs must be equal. The eye therefore has no means to distinguish between these two light mixtures and they are said to match.

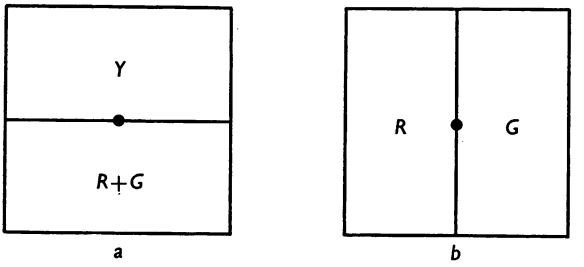

Fig. 6. Two divided colour field sused in a demonstration. *a*, Red + green mixture adjusted to look exactly the same as the yellow above. *b*, Bright red and green used for differential adaptation.

The importance of matching in the analysis of colour vision is this, that when conditions change so that lights which once matched no longer do so, it must be the cone pigments that have changed, it cannot be the nerve processing. This is so important a fact and so often misapprehended that it is worth illustrating it by two examples (in the Lecture these were demonstrated).

Example 1. A field is presented as shown in Fig. 6*a*, where in the lower half, a red + green mixture is so adjusted in proportion that the colour is the same as the yellow in the upper half; the yellow is then adjusted in brightness so that the two halves become identical. This is a famous matching condition (the Rayleigh equation). The B cones are not excited by either field and the quantum catch in R cones is the same from the (Y) and from the ($R+G$) fields; similarly for G cones. The eye is now turned to the bright red/green field of Fig. 6*b*, and fixates for 15 sec upon the centre of that Figure. This procedure in no way affects the absorption spectrum of the cone pigments, but it does affect the relative red/green *adaptation*, i.e. the processing of the nerve signals from the two regions of the retina, so when the adapted eye now observes Fig. 6*a* again, and fixates upon its centre, instead of seeing the field as a uniform yellow as before, it is now decidely reddish on the right and greenish on the left. But each of these vertical reddish or greenish stripes is the same colour above and below. Thus, comparing conditions left and right across the field (Fig. 6*a*), the lights are identical but the appearance quite different; up and down across the field, however, the appearance is identical but the lights are quite

different. This is exactly what was to be expected. The quantum catch is always the same from (Y) and from ($R+G$), thus the input must always be the same and hence the output must always be the same. And though change of nerve processing can change very greatly the colour appearance, it must change both equally since the processing applies to both equally.

Example 2. In contrast to this, replace the normal observer by the commonest type of colour defective (a deuteranomalous trichromat). He will disagree with the normal setting of Fig. 6*a* as Rayleigh (1881) first showed, and change the red + green mixture so that he matches yellow with a colour which to normal eyes is quite green. It is often incorrectly said that this subject simply has less of the normal green pigment or that for some other reason his 'green' nerve signals are smaller than normal. But this would no more cause a change in match than did adaptation to strong green light in the former example. If the anaomalous subject has normal pigments, the chance of each of those molecules catching quanta from (Y) and from ($G+R$) will be normal, and hence he will accept the normal match. The fact that he rejects it must mean that he has at least one pigment abnormal.

In fact deuteranomalous subjects have normal R and B cone pigments, protanomalous have normal G and B cone pigments (Von Kries, 1897; Mitchell & Rushton, 1971*a*, *b*) and their abnormal pigment is still a mystery.

Cone pigment triangle

If we wish to understand more of the relations between the lights, the quantum catches and the mixtures which match, a graphical approach is convenient.

If any light I results in quantum catches, r, g, b in the three cone pigments R, G, B we may represent I by the point on a three-dimensional graph (Fig. 7) with axes OR, OG, OB mutually at right angles. We obtain the point I by moving r units along the R axis, g units upwards and b units out from the paper parallel to the B axis. If the light I were to be increased in energy n times without change in spectral composition, then r, g, b would all be increased n times and I would move to I_1 on the same radius vector, n times the distance from O. Now when a light energy increases without spectral change we say, 'It has got brighter but is still the same colour'. Thus points on the same vector from O represent different brightnesses of the same colour; but vectors in different directions represent different $r:g:b$ ratios and hence different colours. Consequently physiology and psychology have a separate geometry of colour; the physiological input is the quantum catch in the three cone pigments and is represented by the rectangular co-ordinates RGB of Fig. 7. But the nature of nerve

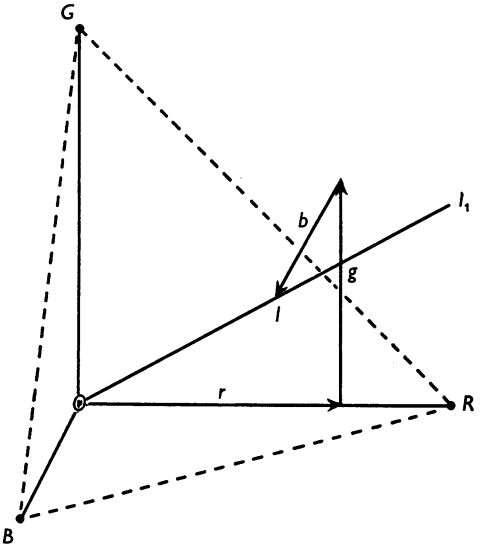

Fig. 7. Three-dimensional plot of the quantum catch in *R*, *G*, *B*, the three pigments. *OR*, *OG* are axes in the plane of the paper, *OB* is perpendicular towards us. The quantum catches *r*, *g*, *b* in *R*, *G*, *B* are all represented by the point *I* in space which is reached by going *r* to the right, *g* up and *b* out of the paper.

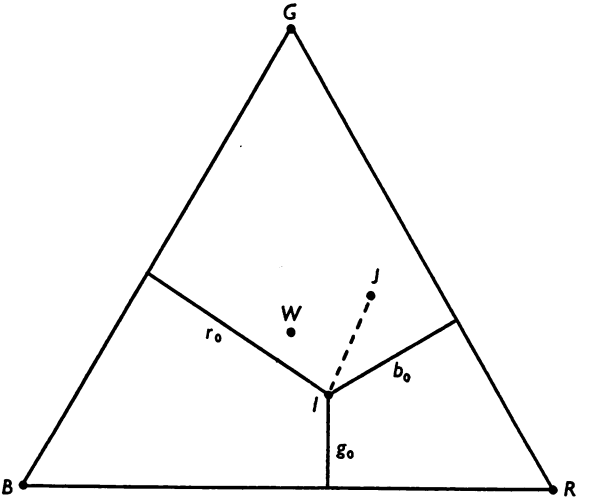

Fig. 8. The triangle *RGB* is that shown by dashed lines in Fig. 7. It is formed by sawing through the solid so as to pass through one unit on each axis. The line *OI*, Fig. 7, pierces this triangle at the point *I*, Fig. 8, and the perpendiculars r_0, g_0, b_0 are proportional to the co-ordinate lengths *r*, *g*, *b* of Fig. 7. If any light of colour *I* is mixed with one of colour *J*, the mixture will have a colour that lies on the line *IJ*.

processing is to transform the outputs into brightness and colour; this is displayed by the polar co-ordinate system of Fig. 7 with brightness represented by the radial distance from O, and colour by the direction of the vector within the solid angle RGB centred on O.

Consider the plane through RGB, points at unit distance from O (Fig. 7). It will be pierced by the vectors corresponding to every colour. Fig. 8 represents this plane and shows RGB as a triangle in which every colour will have a place as defined by the point pierced by its vector, the polar correspondent of the $r:g:b$ proportions of the relative quantum catches. We may therefore call this the *cone pigment triangle*. If r, g, b are expressed in units such that one of each makes white light, then white is situated in the centre of the triangle, Fig. 8. From the way that in Fig. 7 the angle of the polar vector OI is formed from the quantum catches r, g, b it follows that the point I Fig. 8 where it pierces the plane of the triangles has the following property: that the perpendiculars r_0, g_0, b_0 from I on to the sides opposite R, G, B are proportional to the quantum catches r, g, b. In fact if the triangle is of height h

$$\frac{r_0}{r} = \frac{g_0}{g} = \frac{b_0}{b} = \frac{h}{r+g+b}. \tag{2}$$

The effect of mixing two lights I and J (Fig. 8) is found simply by adding the two vectors of Fig. 7. By the parallelogram rule, the resultant lies in the plane of the added vectors and must pierce the triangle somewhere on the line IJ. It will be nearer J the greater the J vector, as follows exactly from the parallelogram rule of vector addition. This is the most useful feature of the colour triangle. Take, for instance, complementary colours which are those which, mixed in suitable proportions, make white. They must therefore both lie on a straight line through the white point, W, one on either side, and all the points so situated on all the lines through W must be complementary.

It is important to know where in the pigment triangle lie the points corresponding to spectral lights, *the spectral locus*. This can be found from Fig. 5 where the R cones' total quantum catch from an equal energy white light will be a simplification of the general expression, eqn. (1)

$$\text{total quantum catch} = n_1 r_1 + n_2 r_2 + n_3 r_3 + \ldots \tag{1}$$

for, ordinates are added at every wave-length and all the n's are equal (for equal energy white). The sum is therefore the area under the R curve, and since this area is scaled to be the same as that under G and B, white light will produce equal catches in all three pigments. This was the scale of units adopted to bring W into the centre of the cone pigment triangle (Fig. 8) and consequently an ordinate drawn at any wave-length λ_1, in Fig. 5

giving ordinates r_1, g_1, b_1 quantum catches, may be transformed into r_0, g_0, b_0 perpendiculars (Fig. 8) according to eqn. (2). The curve in Fig. 9 shows the spectral locus obtained in this way for a set of λ values marked in (nm) upon it. It is seen that though the locus runs close to the *R* and *B* corners of the triangles it keeps its distance from *G*. This asymmetry has implications which have sometimes puzzled people, but it is what would be expected from the curves of Fig. 5. In the far red, the *R* curve is responsible for most of the quanta caught, in the far blue and *B* curve likewise. That means in Fig. 9 that at these extremities the spectral locus is

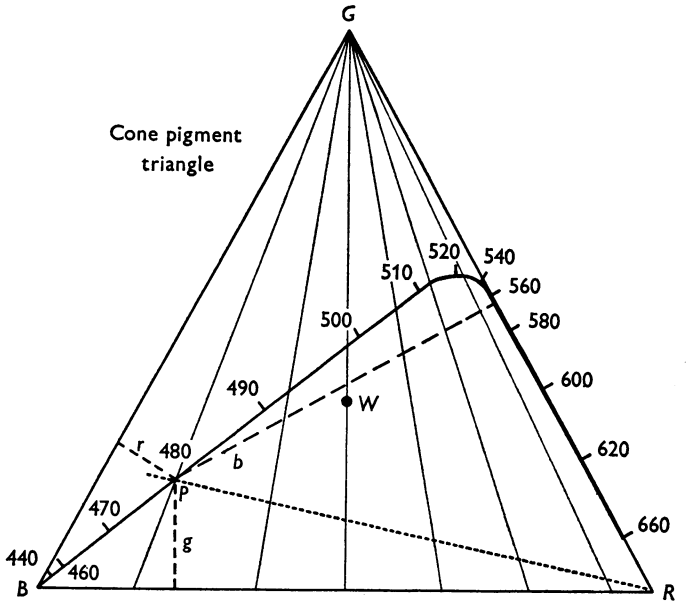

Fig. 9. The cone pigment triangle Fig. 8 with white at centre is without any arbitrariness. Fig. 9 shows the position in it of pure spectral lights of wave-length indicated (nm). The red end of the spectrum approaches *R* along the *GR* side, but does not quite reach *R*. The violet end approaches *B* with a final downward curve, and does not quite reach *B*. The protanope (red-blind) confuses any colour *P* with all the colours which a normal subject will place on the dotted line *R*, *P*. The deuteranope (the other common dichromate) confuses *P* with all those on *GP*. The tritanope (blue-blind) confuses *P* with those on *BP*.

far from the side *BG*, *RG* respectively. But in Fig. 5 there is no wave-length where *G* has a catch which much exceeds *both* the others' and so the locus in Fig. 8 hardly rises more than half way to *G*, and in practice it is impossible to bleach away the green pigment without also bleaching substantially at least one of the others.

The dichromats confusion lines

Clerk Maxwell (1890) investigated the colours which a protanope confuses with any colour P (Fig. 9) and plotted them on his colour triangle (of which Fig. 9 is a projection). He found that all colours confused with P lay on the straight line joining P to a point just outside the red corner of his triangle. They lie in Fig. 9 exactly on the line PR. As Maxwell pointed out, this type of confusion means that the red dimension of colour is absent in protanopes. Fig. 9 interprets this in terms of quantum catches. In Fig. 9 the R corner is the place where the R cone pigment alone is excited, G and B having zero perpendiculars. The dotted line PR represents the range of colours resulting from the mixture of P and R in various proportions. Since to the protanope this range is all identical in colour, a change in the R quantum catch is without effect. Thus the R cone pigment must be either absent or its quantum catch does not contribute to vision. In fact, as we have seen, erythrolabe is absent. In the pigment triangle, Fig. 9, the colours which the protanope confuses with any colour P are found to lie on the line PR. Those which the deuteranope confuses with P are found to lie on PG. Those confused by the tritanope are found to lie on PB. Thus for each kind of dichromat the lines of confusion are seen to radiate from the corner of the pigment he lacks; this is shown in Fig. 9 for the deuteranope, where they radiate from G. It follows that the spectral wave-length which to a dichromate appears white is the point where the line joining W to the radiating corner cuts the spectral locus, e.g. 499 nm for deuteranopes and 495 nm for protanopes.

Koenig primaries

In the foregoing account we have touched on two independent ways of finding the spectral sensitivity of the cone pigments. The most straightforward to understand is the micro-spectrophotometer results of Fig. 1 and we have proceeded as though all results could be deduced from those – as indeed they could if it were not for some divergence in the experimental points and the many difficult small correction factors that have to be applied, into which we shall not enter.

In particular from Fig. 1 we could deduce the colour matching functions – the energies of three chosen spectral lights which when mixed match perfectly unit energy of every monochromatic light. Wright (1929) measured these energies so accurately that when Stiles (1955) repeated and extended them for the International Commission (C.I.E.) he confirmed Wright almost to perfection. This accuracy is much better than can be achieved by cone densitometry so it is clearly desirable to proceed from colour matches to the pigment curves rather than in the opposite direction.

Indeed for a century attempts have been made to do just this. But, as is well known, it is easier to mix than to unmix, and we need a little more information besides colour matching to unmix the spectral colour mixture functions. Koenig pointed out that dichromats have confusion lines (Fig. 9) as though they each lacked one normal pigment, and if this were true we have from the lie of those lines the information necessary to transform from colour matching to separate cone sensitivities. The results of densitometry on the eye of protanopes and deuteranopes, as we have seen, prove objectively that dichromats do indeed lack one pigment, and thus justify Koenig's transformation. From it we may obtain Fig. 5 and hence Fig. 9 using only the very exact C.I.E. measurements of the spectral colour mixtures, and the well defined concurrence points of the confusion lines in the three types of dichromacy. And that is actually the way that those Figures were plotted (following Wyszecki & Stiles, 1967), not from the considerably less accurate measurements of Fig. 1.

The colour triangle recommended by the International Committee, and hence called the C.I.E. triangle is a different representation of these same experimental results and hence has exactly the same accuracy as the cone pigment triangle (Fig. 9). Indeed the two triangles would be simply mathematical projections one of the other except that in Fig. 9 light energies are expressed not in ergs but in quanta. Only in these units will the spectral sensitivity coincide with the pigment's absorption spectrum. With the International Committee behind it, it is the C.I.E. triangle that is given in all the best books and nearly everywhere when colour is to be geometrically represented. But I shall not give it here, for I think the cone pigment triangle (Fig. 9) vastly superior as an indication of what is going on in the eye. The C.I.E. committee are colour physicists concerned not with the mechanism of vision but with the application of colour measurements to technology. The C.I.E. triangle is brilliantly ingenious as an aid to the calculations of chromaticities which can be upheld in a court of law where colour specification is in dispute. But that triangle is monstrous as an indication of what is going on in the mechanism of vision. It displays all colours as a mixture of three primary lights, none of which have an existence that can be easily imagined. One of the three primaries is bright; it is a pure green from which is subtracted a lot of red which it does not contain. The other two primaries are quite dark; they have strong colour but zero luminance. These do not seem to me ingredients that lead to clarity in our conception of colour mechanisms and I am astonished that some physiologists and many psychologists employ them to instruct the young and bewilder the old.

The input to the eye is certainly the quantum catch in the receptors, and colour is largely defined by the relative catches in the three kinds of cone.

The cone pigment triangle Fig. 9 displays this relation rather clearly; and the way that the triangle is related to the three dimensional graph of Fig. 7 permits an easy mental switch from the one to the other. The physiological process is best displayed by the cartesian catch r, g, b of the three independent cone pigments, the psychological transformation by the polar representation, where $\mathbf{r}$, the distance from O, indicates brightness and where colour is located by the point where the vector $\mathbf{r}$ pierces the cone pigment triangle.

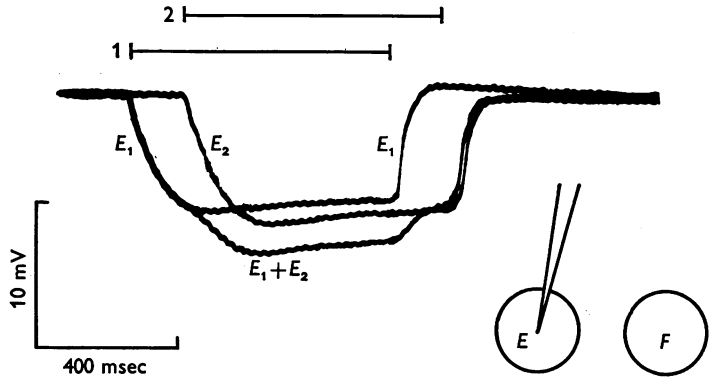

Fig. 10. Amplitude-variant intracellular potentials in response to light. Horizontal lines 1, 2 show time course of lights shone on to micro-electrode area E either singly or together. E_1 is response to light 1 alone, E_2 to 2 alone, E_1+E_2 to both together. The response follows the time course of the light but the potential is not linear with light intensity. Records are from horizontal cells (L type) in the fish (*Tinca*). Note the response to light is hyperpolarization (from Naka & Rushton, 1967).

Part II. Nerve signals

The manner in which signals are transmitted by long nerves such as the sciatic of the frog or the giant axon of the squid is very well known. When the cell membrane is sufficiently depolarized it suddenly and momentarily becomes permeable to Na ions which rush in and cause depolarization of the neighbouring segment, and so the activity spreads rapidly down the nerve. The process is self-limiting both in duration and magnitude and results in the familar all-or-none spike of potential which is followed by a refractory period during which the nerve is inexcitable.

In 1953, Svaetichin with very fine electrodes obtained from the retina intracellular records which behaved in a totally different manner. The potential change in response to light is with some cells a depolarization, but more commonly a hyperpolarization. It goes on with the light, stays on with the light and goes off with the light (Fig. 10). In this respect therefore the potential change is like that from a photocell. The size of potential

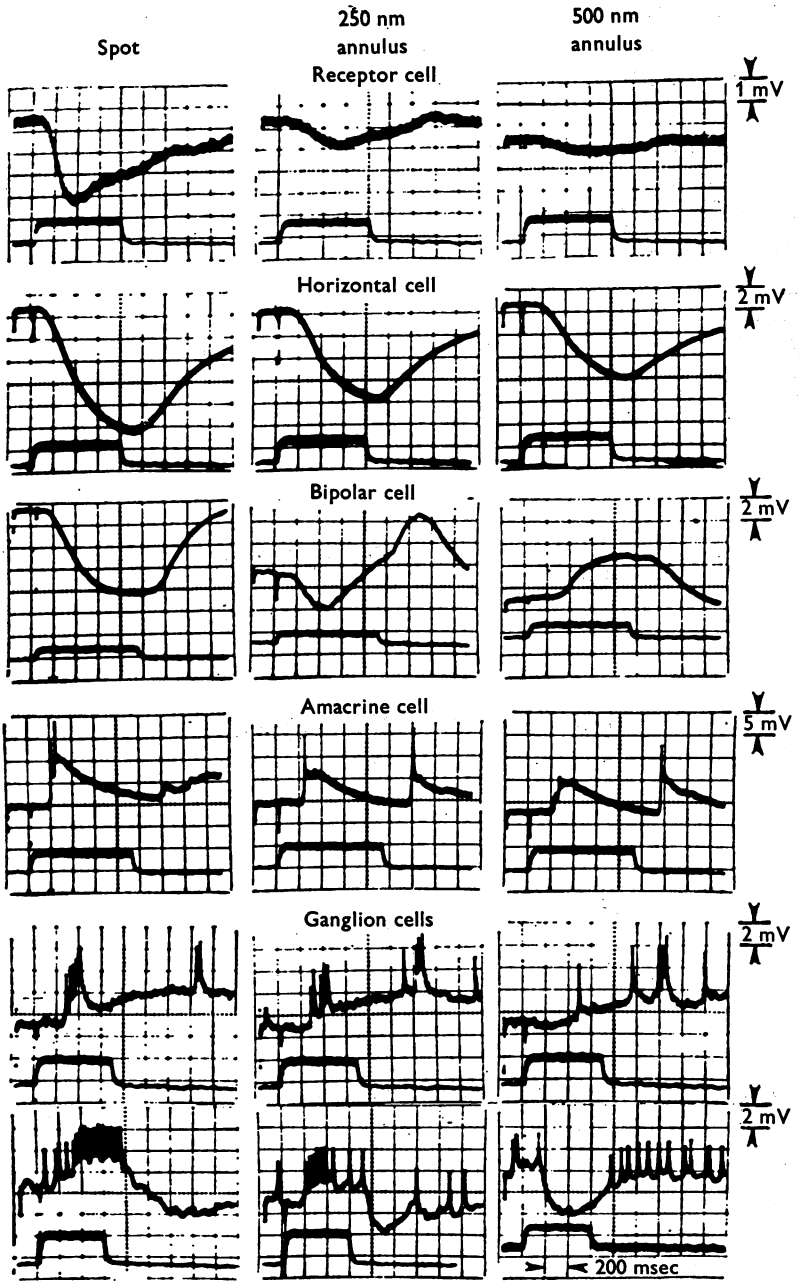

Fig. 11. Intracellular responses recorded at all levels in the (large-celled) retina of the mud-puppy (a larval amphibian form). Only ganglion cells give the familiar train of all-or-none nerve spikes. Amacrines seem to exhibit one transient spike. Earlier in the transmission line only amplitude-variant messages are recorded (from Werblin & Dowling, 1969).

change also increases with light intensity but not linearly. In Fig. 10 light 1 applied to spot E gave the potential wave (E_1) and light 2 gave E_2. The sum of lights $1+2$ gave potential (E_1+E_2). This is much smaller than the sum of potentials E_1+E_2. It is obvious that this kind of photocell is far more useful for giving time–intensity records of the light than one which can only discharge a series of unit clicks, but the clicks are much more reliable for long-distance transmission. So when the complex retinal interactions have been accomplished in *amplitude-variant* signals, the more important results are encoded in the *frequency-variant* spike signals of the optic nerve for safe sending to the brain.

Amplitude variation is the mode in receptors, horizontal cells and bipolars (Fig. 11). It is not until the signals reach the amacrine cells that

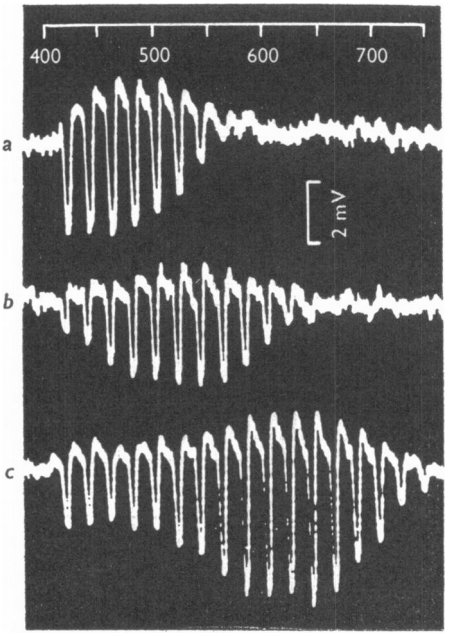

Fig. 12. Intracellular records from inner segments of single cones in the carp. Amplitude of response recorded as a function not of time, but of wave-length. Interposed between light and receptor is a wheel perforated by a ring of windows so that when the wheel is spun the light passes through all the windows in succession. Each window is fitted with an interference filter and a neutral density to pass always the same light energy whatever the wave-length. A potentiometer is coupled to the wheel so that the horizontal shift in the Figure is proportional to the wave-length of light passing through the window. Thus record *a* of this Figure shows the equal energy response of a cone most sensitive in the blue. Records *b* and *c* show maximum sensitivity in green and red respectively. All responses to light are hyperpolarizing (from Tomita, Kaneko, Murakami & Pautler, 1967).

spikes are seen (Werblin & Dowling, 1969). Rods and cones all hyperpolarize upon illumination and different cones have different spectral sensitivities (Fig. 12). Hyperpolarization is associated with increased membrane resistance and a diminished resting dark current (Penn & Hagins, 1969). Bipolar cells start the system of 'push-pull' that is found at many levels in the visual pathway. These are paired nerves responding in opposite directions, the one hyperpolarizing the other depolarizing, the one responding when the light goes on, the other when it goes off, the one sensitive to illumination at the centre of the 'receptive field', the other at the periphery, the one responsive to red, the other to green, etc. These contrasts do not all occur together in one pair of nerves, but where features are to be contrasted there seems to be a pair of rather symmetrical nerves to emphasize the difference.

The direct path from receptors to brain is via bipolar cells and ganglion cells whose axons form the optic nerve. Spreading through the layers of the retina at right angles to those radial communication lines are horizontal cells and amacrine cells.

The horizontal cells lie at the junction between the receptors and the bipolars and send processes to those junctions. It seems likely that the function of these transverse connexions is not to spread laterally the pattern of excitation and so smudge the picture, but rather to carry out an automatic gain control, which will allow the eye to discriminate over a wide range of luminosities. An automatic camera has a photocell which measures the mean brightness of the scene and adjusts the stop so that the mean intensity of light falling on the film lies in the middle of its sensitivity range. In a similar manner one horizontal cell receives receptor signals over about 0·5 mm of retinal surface and probably scales down the signals in all the bipolars in that region so that the mean level is maintained nicely in the bipolar working range, and at the same time the full contrast of the picture is preserved. The well known Weber–Fechner law, the way that the threshold for discrimination rises in proportion to the background luminance is an expression of this organization where contrast is preserved independent of light level.

The earliest amplitude-variant potentials were recorded by Svaetichin (1953) from horizontal cells, and these records are often called S-potentials. They are of two kinds: (*a*) those that always hyperpolarize in response to light as in Fig. 10 and Fig. 11, which are called *L* (= luminosity) potentials, and those (*b*) which hyperpolarizes to blue light and depolarize to red, and are called *C* (= colour) potentials. Fig. 13 shows an example from the horizontal cells of fish, which have 'double cones', a red and green cone in close apposition so that current flowing between them to hyperpolarize the green will depolarize the red. Perhaps the *C* potentials result entirely

from the responses of double cones, which are not present in the retina of man.

Fig. 11 shows that the amacrine cells respond chiefly to a change in light reacting by sudden depolarization both to 'on' and 'off'. Only ganglion cells respond to steady light by a steady stream of impulses.

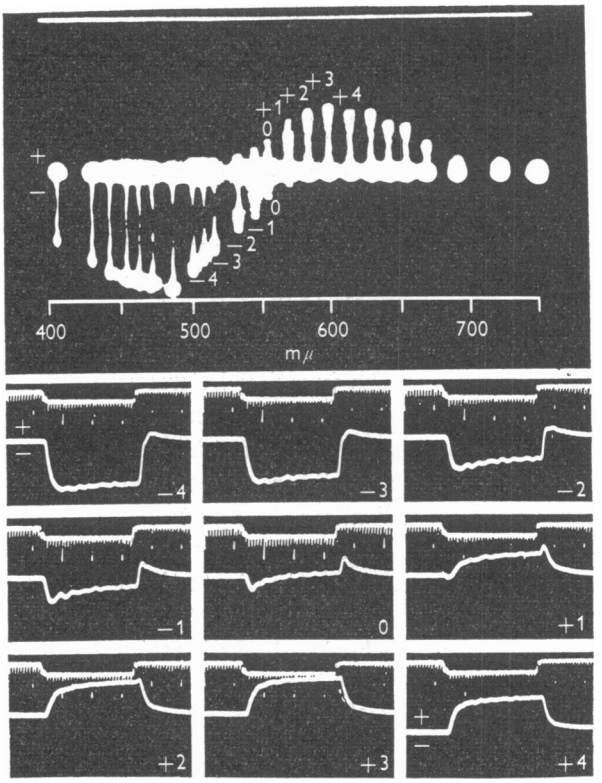

Fig. 13. Intracellular S-potentials from horizontal cells in fish (*Mugilidae*); upper record, amplitude recorded as function of wave-length by the spinning-wheel technique used in Fig. 12. But here only short wave-lengths hyperpolarize, long wave-lengths of light depolarize. The time course of responses −4 to +4 are displayed in the lower records where the scale is in tenths and hundredths of a second (from MacNichol & Svaetichin, 1958).

The difficulty at present in obtaining a clear view of the retinal mechanisms is (*a*) that these in detail must be extremely complex, as the structure of the retina shows, and consequently simple interpretations will almost certainly be wrong; (*b*) that the necessary intracellular recording demands the highest technical skill even to obtain a few brief records, and prolonged analysis of any single cell is often out of the question; (*c*) only

comparatively large cells in cold blooded excised eyes are suitable for much of the work and their organization may be very different from that in human eyes; (*d*) it is one thing to obtain a potential record and a very different thing to say what part it plays in vision.

In what follows I shall be frankly speculative and try to give an over-simplified view of the transmission of information in the amplitude-variant system. Detailed records of high quality pour into the literature, and the deficiencies of this scheme will doubtless soon be apparent enough, if not from the start.

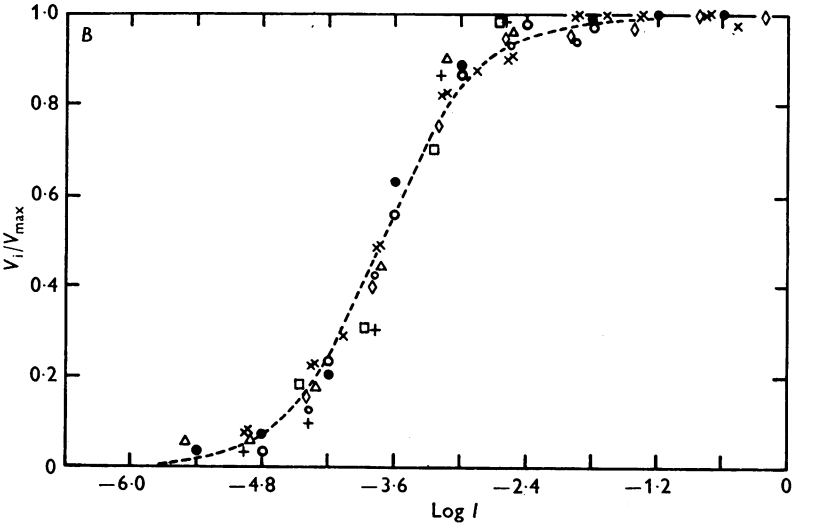

Fig. 14. Intracellular responses of ten turtle cones to light flashes of various intensities I. The responses have been scaled to the same maximum and the curve shifted laterally to normalize the sensitivity. The dashed curve is an H_1 curve (from Baylor & Fuortes, 1970).

Signal as a function of light intensity

Svaetichin (1953) showed for his S-potentials that the potential V generated by a light pulse of intensity I was linear with log I over a range of some 2 log units. Naka & Rushton (1966, 1967) found that when only a single class of cone is involved, the relation

$$V/V_\infty = I/I+\sigma) = H(I) \qquad (3)$$

is a better description and fits the whole range. When I becomes very large V saturates at V_∞; when $I = \sigma$, V half-saturates since $V/V_\infty = \frac{1}{2}$.

This hyperbolic relation between V and I seems to hold rather widely in the visual system (including invertebrates) and it is useful to have a label for it. We may call V plotted against I an 'H curve'; V plotted against

log I an 'H_1 curve'; and the double log plot, log V against log I an 'H_2 curve'.

The record of Fig. 14 is taken from Baylor & Fuortes' (1970) measurements of intracellular potential in single cones of the turtle. The points measure the peak potential resulting from flashes of light whose log energy is shown as abscissa. The dashed curve they have drawn through the points is the H_1 relation. The same relation is found in the potentials from horizontal cells (Naka & Rushton, 1967) and a similar relation seems to hold also for bipolars. It looks as though this may be a characteristic in the transmission of amplitude-variant signals.

This would follow if a synaptic transmitter, released in amount x, generates a post-synaptic potential of magnitude $V_1 = H(x)$. If now this potential spread across the cell and at the other end released a transmitter y in amount proportional to V_1, we might expect y to generate a post-synaptic potential V_2 that was $H(y)$. It is a simple piece of algebra to show from these two relations that V_2 is also an H function of x (with a different σ value), and therefore in a long train of cells operating in this simple way there will be an H function between input and V_n the potential recorded at every stage. Now from Fig. 14 it is clear that the receptor potentials are certainly H functions of light intensity, and so we might expect all the potentials down the line to be H functions of light intensity so long as amplitude-variation holds, i.e. up to the amacrines and ganglion cells. And indeed this seems to be the case.

The human eye

The intracellular records that we have been considering have two disadvantages: (*a*) they are obtained from eyes organized very differently from ours and (*b*) the range of signal strength measurable is less than 2 log units. It is therefore of interest that by use of a special technique, Alpern, Rushton & Torii (1970*a*, *b*, *c*) were able to show an H relation between signal strength and light intensity in man all the way from threshold up to saturation some 100,000 times more intense.

Spekreijse & van den Berg (1971) have shown by records from the ganglion cells in the goldfish that a linear interaction occurs in spatial and colour signals which can be analysed by using the ganglion cell as a *null* output and thereby disengaging from the non-linearities of spike encoding. Donner & Rushton (1959) and Rushton (1959) did somewhat the same in the frog. In man Alpern *et al.* used the same device balancing a test flash against an inhibitory annular surround flash so that the encoded output (threshold for seeing the flash) remained invariant. Former work (Alpern & Rushton, 1965) had shown that this interaction was receptor-specific; the rods of the test flash were inhibited only by rods in the surround, no matter how

strongly cones were also excited; red cones of the flash were inhibited only by red cones in the surround, etc. The later work studied both for rods and for cones the size of nerve signal as a function of light intensity, and also the effect upon signal size of a steady background. The results came out extremely simple. If N is the size of nerve signal, ϕ is the intensity of test

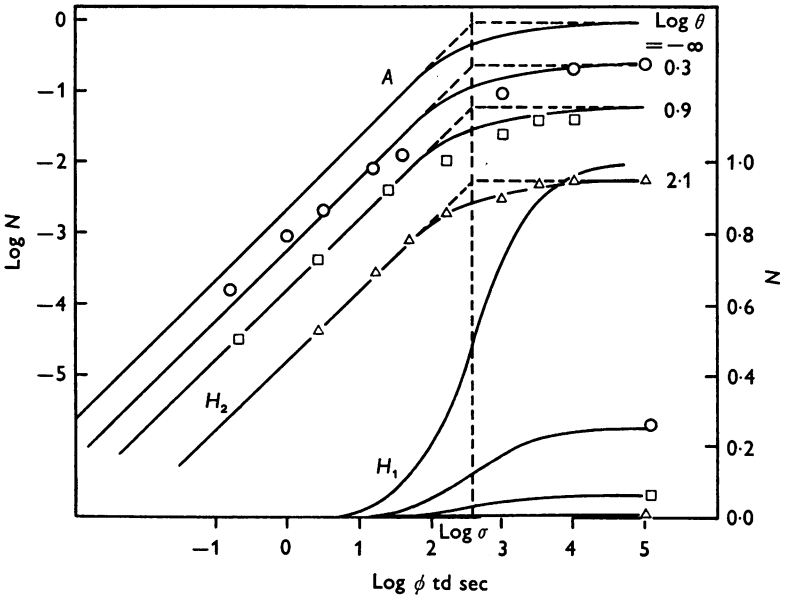

Fig. 15. Size of human nerve signals N as a function of ϕ the exciting flash and θ the steady background. Lower curves H_1 plot signal size (scale on right) against log ϕ. The curves are exactly like those of Fig. 14, and the scaling factor is proportional to the steady background θ. The upper curves plot log N (scale on left) and show the H_2 curves which fit the experimental points (including the top curve where the points are not shown) over a range of 4 log units (from Alpern, Rushton & Torii, 1970*b*).

flash (measured as the quantum catch in the particular receptors investigated, e.g. the chlorolabe cones) and θ the intensity of steady background upon which ϕ falls (also measured as quantum catch in chlorolabe cones) then

$$N = \frac{\phi}{\phi+\sigma} \cdot \frac{\theta_D}{\theta_D+\theta}. \tag{4}$$

The constant θ_D is the receptor noise or *eigengrau*. When there is no luminous background eqn. (4) reduces to the form of eqn. (3). The effect of a background is to attenuate the size of signal in proportion to $(\theta_D+\theta)$.

Fig. 15 from Alpern *et al.* (1970*b*) shows for rods some experimental results with curves which are all described by eqn. (4). The upper curves are H_2 where log N is plotted against log ϕ for various values of log θ.

The lower curves are the same results displayed as H_1 where N (not log N) is plotted against log ϕ, as is common with the results of electrophysiology. Electrophysiology can seldom record below 1 % of the maximum (saturated) potential, that is below -2 on our H_2 scale (on the left), but our human thresholds are recorded accurately another 2 log units below this and continue to fit the H_2 curve.

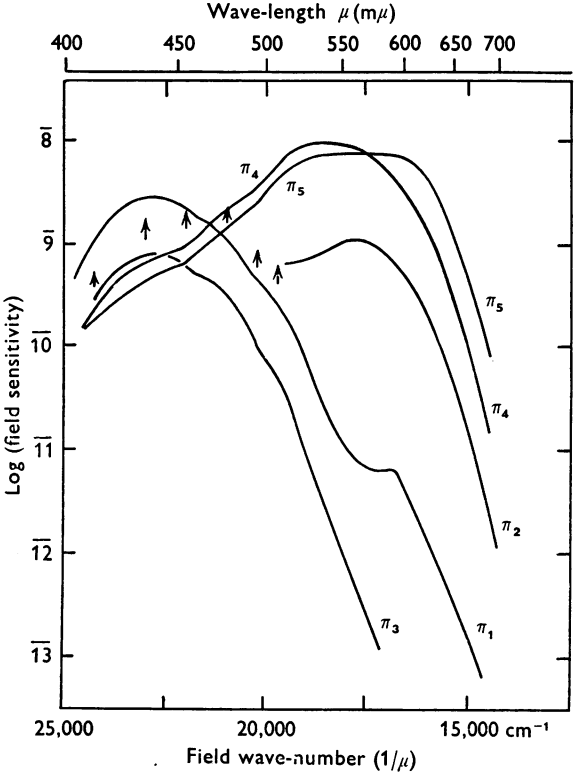

Fig. 16. Spectral sensitivity of Stiles's cone mechanisms determined by his two-colour increment threshold technique. To a good approximation π_5, π_4, π_3 correspond to the sensitivity of the R, G, B cone pigments. The hump in π_1 and the nature of π_2 are still not known (from Stiles, 1953).

Cones fit this same pattern with σ, the semi-saturation constant and θ_D, the receptor noise both 100 times as great as with rods. For both rods and cones the threshold is when $N = 10^{-5}$, thus over much of the range of ordinary vision eqn. (4) reduces to

$$N = \frac{\theta_D}{\sigma} \cdot \frac{\phi}{\theta},$$

i.e. the signal is proportional directly to the flash and inversely to the background.

Two-colour thresholds

An important contribution to our understanding of colour has been made over some 20 years by Stiles (1939, 1946, 1949, 1953, 1959). With great accuracy he has measured the threshold for the detection of a flash of one wave-length when seen superimposed on a steady background of another. To a first approximation his results are as follows. There are four principal mechanisms, rods and three cones whose spectral sensitivities correspond to that of the three cone pigments erythrolabe, chlorolabe and cyanolabe; π_5, π_4, π_3 of Fig. 16. Each mechanism is independent of the others and is excited by ϕ, the quanta it catches from the flash and is each depressed by θ, the quanta it catches from the steady background. The relation he finds between ϕ and θ at threshold is

$$\phi = k(\theta + \theta_D). \tag{5}$$

The threshold for seeing was determined by whichever mechanism had the lowest threshold in the given circumstance.

This relation, which fits such a wide array of observations, accords exactly with eqn. (4). When the N signal is at threshold it is a constant number of about 10^{-5} and thus ϕ is negligible compared with σ. Therefore

$$\phi = \frac{\sigma N}{\theta_D}\left(\theta + \theta_D\right) \tag{6}$$

which is the same as Stiles' relation (5) with the components of k defined.

In one important paper (Anguilar & Stiles, 1954) the background fields for rods were pushed up to the point where rods saturated. In that condition ϕ is no longer negligible compared to σ and in fact their results fit our eqn. (4) over the whole range and constitute a particular case of the more general conditions of saturation that we studied (Alpern *et al.* 1970*b*).

Colour signals

I do not think that any of the signals we have been considering are specially organized for colour. They originate of course in colour-sensitive cones and are attenuated probably by the horizontal cells which constitute a gain control of strictly the same colour sensitivity. This, if it applied individually to every cone signal, would go a long way to abolish colour and most other features from visual signals. But the attenuation is not individual. All the members of (say) the red cone local community are attenuated by a factor proportional to their average output. Thus the mean output remains constant but the contrast between the individual outputs of Jack and his neighbour remains stark. If this is done for all three classes of cone, the result will be that the coloured appearance of

objects will not change very much with different colours of illumination, so colour and brightness-dependent reflexes learnt at noon will not have to be relearnt near sun-down. The two-colour thresholds of Stiles are really brightness thresholds of a single class of cones, and whether the background looks the same colour as the flash or not makes no difference to the threshold so long as the quantum catch in those cones is the same.

We have accepted the suggestion derived from Young that colour depends upon the relative quantum catch in the three cone pigments, but have said nothing about who computes what the relative catch is. There seem in fact to be several different counting houses balancing in pairs sometimes red against green, yellow against blue, brightness against darkness, centre against periphery, etc. In the monkey, cells have been found in the lateral geniculate which respond specifically to changes of colour (De Valois, 1965). Wiesel & Hubel (1966) studied that organization in relation to the receptive fields and found most of the cells to have a centre-surround organization of the field in the retina with opposing colours, e.g. red centre versus green surround. Almost any combination of colours and on/off organization were found. As seen from Fig. 1 a green light is bound to produce a large quantum catch in the red cones, so in any region where both cones are present it cannot be supposed that green light stimulates only (or even mainly) green cones. Red cones seem more plentiful (see Fig. 2).

No doubt it is upon the output of red/green opponent cells and similar cells for other cone input combinations that our sense of colour depends, and side-by-side colour contrast must result (at least in part) from this centre-surround colour organization. But unfortunately simultaneous and successive contrast are a great deal easier to describe than to analyse. They are closely bound with the continual eye trembling which is a feature of vision so important that when it is stopped (as in a stabilized image) all colour and most detail fades away in a matter of seconds.

The fact is that colour, like brightness, is only a *comparative* measure, a balance of paired inputs from different kinds of cone in different localities fluctuating with eye tremor, and this is what makes our judgements so relative. The surprising thing is that despite their continual deceitfulness we persist in believing our eyes. In order to do this we have to strain the imagination somewhat. That is an aspect of colour adaptation quite different from those we have so far considered, and I shall end by describing one example which impressed me. It was seen with Land's celebrated two-colour projection technique (Land, 1959).

Two photographs are taken of some coloured scene, one through a filter transmitting all lights of wave-length greater than Na light, the other through a filter transmitting all lights of wave-length shorter than Na

light, and a lantern slide black/white (positive) is made from each. These two slides are projected in register on to a screen so that all the contours coincide exactly. A red filter is placed in the beam projecting the long wave slide and a neutral filter is placed in the other beam to make the brightness of both about the same. A pleasing picture is seen in which some region may appear as a very good green.

It is clear that this green region (like every other) is a mixture of white light and a red, which could be a monochromatic light of 620 nm without essentially changing the appearance. You may say 'This is easily explained. There was little or no red on the "green" region, there was much more all around; the patch looked green by simultaneous contrast.' Very good. Now comes the interesting observation. I looked at the 'green' region through a tube narrow enough to cut out all the surround. Interposing the tube or removing it made no difference at all to the 'green' colour. I then asked a colleague to block the red projector while I still looked steadily down my tube. I could just see when he did it, for the red had sent a weak light onto the 'green' patch so when it was blocked the patch went *greener*. And as I continued to watch through my tube nothing further happened. I was, of course, now looking at the black/white/grey picture with only the neutral filter interposed in the projection from a black/white slide, and I knew it. But the patch down the tube was a good green to me – well, it was certainly a good green at the start, and nothing further had happened so it must still be a good green mustn't it? The green did not fade, but my conviction faded that it was green. I began to ask myself how I knew that it was green, and could not answer. But neither can I answer that when I look at fresh grass. I removed the tube for an instant and replaced it, but that momentary glance upon the cold grey picture utterly froze the green from the end of my tube and I could not begin to recapture it (Rushton, 1961). It seems most unlikely that any sudden change in *retinal* organization accompanied that instant glance, but my colour concept changed instantly and I woke to the grey reality as from a dream. But, as I said at the beginning, the content of our hollow scientific structures *is* such stuff as dreams are made on: there is nothing either green or grey but thinking makes it so.

It is a pleasure to thank the Medical Research Council for their support.

REFERENCES

AGUILAR, M. & STILES, W. S. (1954). Saturation of the rod mechanism of the retina at high levels of stimulation. *Optica Acta* **1**, 59–65.

ALPERN, M., LEE, G. B., MAASEIDVAAG, F. & MILLER, S. S. (1971). Colour vision in blue-cone 'monochromacy'. *J. Physiol.* **212**, 211–233.

ALPERN, M. & RUSHTON, W. A. H. (1965). The specificity of the cone interaction in the after-flash effect. *J. Physiol.* **176**, 473–482.

ALPERN, M., RUSHTON, W. A. H. & TORII, S. (1970*a*). The size of rod signals. *J. Physiol.* **206**, 193–208.

ALPERN, M., RUSHTON, W. A. H. & TORII, S. (1970*b*). The attenuation of rod signals by backgrounds. *J. Physiol.* **206**, 209–227.

ALPERN, M., RUSHTON, W. A. H. & TORII, S. (1970*c*). Signals from cones. *J. Physiol.* **207**, 463–475.

BAKER, H. D. & RUSHTON, W. A. H. (1965). The red-sensitive pigment in normal cones. *J. Physiol.* **176**, 56–72.

BAYLOR, D. A. & FUORTES, M. G. F. (1970). Electrical responses of single cones in the retina of the turtle. *J. Physiol.* **207**, 77–92.

BOLL, F. (1876). Zur Anatomie und Physiologie der Retina. *Mber. Akad. Wiss. Berlin* **41**, 783–787.

CRESCITELLI, F. & DARTNALL, H. J. A. (1953). Human visual purple. *Nature, Lond.* **172**, 195–197.

DE VALOIS, R. (1965). Analysis and coding of colour vision in the primate visual system. *Cold Spring Harb. Symp. quant. Biol.* **30**, 567–579.

DONNER, K. O. & RUSHTON, W. A. H. (1959). Retinal stimulation by light substitution. *J. Physiol.* **149**, 288–302.

HOOD, C. & RUSHTON, W. A. H. (1971). The Florida retinal densitometer. *J. Physiol.* **217**, 213–229.

KOENIG, A. (1894). Über den menschlichen Sehpurpur und seine Bedeutung für das Sehen. *S.-B. dtsch. Akad. Wiss., Kl. med. Wiss.* 577–598.

KÜHNE, W. (1878). Zur Photochemie der Netzhaut. *Unters. physiol. Inst. Univ. Heidelberg* **1**, 1–14.

LAND, E. H. (1959). Color vision and the natural image. Part I. *Proc. natn. Acad. Sci. U.S.A.* **45**, 115–129.

LOMONOSOV, M. V. (1757). Lecture on the origin of light, presenting a new theory of colours (in Russian). St Petersburg: Academy of Sciences.

MACNICHOL, E. F. & SVAETICHIN, G. (1958). Electric responses from the isolated retinas of fishes. *Am. J. Ophthal.* **46**, 26–40.

MARKS, W. B. (1965). Visual pigments of single goldfish cones. *J. Physiol.* **178**, 14–32.

MARKS, W. B., DOBELLE, W. H. & MACNICHOL, E. F. (1964). Visual pigments of single primate cones. *Science, N.Y.* **143**, 1181–1183.

MAXWELL, J. C. (1890). *Scientific Papers*, vol. 1, pp. 126–154. Cambridge: Cambridge University Press.

MITCHELL, D. E. & RUSHTON, W. A. H. (1972*a*). Visual pigments in dichromats. *Vision Res.* (in the Press).

MITCHELL, D. E. & RUSHTON, W. A. H. (1972*b*). The red/green pigments of normal vision. *Vision Res.* (in the Press).

NAKA, K. I. & RUSHTON, W. A. H. (1966). S-potentials from colour units in the retina of fish (Cyprinidae). *J. Physiol.* **185**, 536–555.

NAKA, K. I. & RUSHTON, W. A. H. (1967). The generation and spread of S-potentials in fish (Cyprinidae). *J. Physiol.* **192**, 437–461.

NEWTON, I. (1672). New theory about light and colours. *Phil. Trans. R. Soc.* **80**, 3075–3087.

PENN, R. D. & HAGINS, W. A. (1969). Signal transmission along retinal rods and the origin of the electroretinographic *a*-wave. *Nature, Lond.* **223**, 201–205.

RAYLEIGH, LORD (1881). Experiments on colour. *Nature, Lond.* **25**, 64–66.

RUSHTON, W. A. H. (1955). Foveal photopigments in normal and colour-blind. *J. Physiol.* **129**, 41–42*P*.

RUSHTON, W. A. H. (1956). The difference spectrum and photosensitivity of rhodopsin in the living human eye. *J. Physiol.* **134**, 11–29.

RUSHTON, W. A. H. (1959). Excitation pools in the frog's retina. *J. Physiol.* **149**, 327–345.

RUSHTON, W. A. H. (1961). The eye, the brain and Land's two-colour projections. *Nature, Lond.* **189**, 440–442.

RUSHTON, W. A. H. (1963). A cone pigment in the protanope. *J. Physiol.* **168**, 345–359.

RUSHTON, W. A. H. (1965*a*). A foveal pigment in the deuteranope. *J. Physiol.* **176**, 24–37.

RUSHTON, W. A. H. (1965*b*). The Newton Lecture: chemical basis of colour vision and colour blindness. *Nature, Lond.* **206**, 1087–1091.

SPEKREIJSE, H. & VAN DEN BERG, T. J. T. P. (1971). Interaction between colour and spatial coded processes converging to retinal ganglion cells in goldfish. *J. Physiol.* **215**, 679–692.

STILES, W. S. (1939). The directional sensitivity of the retina and the spectral sensitivities of the rods and cones. *Proc. R. Soc.* B **127**, 64–105.

STILES, W. S. (1946). Separation of the 'blue' and 'green' mechanisms of foveal vision by measurements of increment thresholds. *Proc. R. Soc.* B **133**, 418–434.

STILES, W. S. (1949). Increment thresholds and the mechanisms of colour vision. *Documenta ophtha.* **3**, 138–163.

STILES, W. S. (1953). Further studies of visual mechanisms by the two-colour threshold technique. *Coloquio sobre problemas opticas de la vision*, pp. 65–103. Madrid:

STILES, W. S. (1955). The basic data of colour matching. In *Physical Society Yearbook* 1955, pp. 44–65. London: Physical Society.

STILES, W. S. (1959). Colour vision: the approach through increment threshold sensitivity. *Proc. natn. Acad. Sci. U.S.A.* **45**, 100–114.

SVAETICHIN, G. (1953). The cone action potential. *Acta physiol. scand.* **29**, suppl. 106, 565–600.

TOMITA, T., KANEKO, A., MURAKAMI, M. & PAUTLER, E. L. (1967). Spectral response curves of single cones in the carp. *Vision Res.* **7**, 519–531.

VON KRIES, J. (1897). Über Farbensysteme. *Z. Psychol. Physiol. Sinnesorg.* **13**, 241–324.

WERBLIN, F. S. & DOWLING, J. E. (1969). Organization of the retina of the mudpuppy *Necturus maculosus* II. Intracellular recording. *J. Neurophysiol.* **32**, 339–355.

WIESEL, T. N. & HUBEL, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophysiol.* **29**, 1115–1156.

WILLMER, E. N. (1946). *Retinal Structure and Colour Vision.* Cambridge: Cambridge University Press.

WILLMER, E. N. (1961). Human colour vision and the perception of blue. *J. theoret. Biol.* **2**, 141–179.

WRIGHT, W. (1929). A re-determination of the trichromatic mixture data. *Spec. Rep. Ser. med. Res. Counc.* **139**, 1–38.

WYSZECKI, G. & STILES, W. S. (1967). *Colour Science.* New York: John Wiley and Sons.

YOUNG, T. (1802). On theory of light and colours. *Phil. Trans. R. Soc.* 12–48.