The Ferrier Lecture, 1980: Critical Limiting Factors in the Design of the Eye and Visual Cortex

H. B. Barlow

THE FERRIER LECTURE, 1980

Critical limiting factors in the design
of the eye and visual cortex

BY H. B. BARLOW, F.R.S.

Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, U.K.

The main factors limiting the performance of the peripheral parts of the visual system can be specified, and doing this clarifies the nature of the interpretive tasks that must be performed by the central parts of the system. It is argued that the critical factor that hinders development of better resolving power is the difficulty of confining light within the waveguide-like outer segment, and that for sensitivity this critical factor is the thermal decomposition of photosensitive pigments. Knowledge of these limits makes many surprising details of the eye intelligible. Under-standing the difficulties posed by the narrow dynamic range of nerve fibres may give similar insight into the coding of the retinal image for transmission to the brain.

Our level of understanding changes when we come to the visual cortex, for although we do not lack good anatomical and neurophysiological data, these do not make the principles of operation self-evident in the way that the structure of the eye immediately suggests that it is an image-forming device. The cortex converts the representation of the visual field that it receives into reliable knowledge of the world around us, and the trouble may be that we lack good models of how this can be done. A system that can respond to single quanta and resolve almost to the diffraction limit is unlikely to employ grossly inefficient methods for those higher functions upon which its whole utility depends, and so it is worth seeking out the limiting factors.

The quality of human performance at certain higher perceptual tasks is high compared with the limit of reliable statistical inference; hence much of the sample of information available in a visual image must be effectively utilized. But there are strong limitations on the connectivity in the cortex, so that one is forced to consider how the relevant informa-tion can be collected together. Three stages of dealing with the visual image are proposed: the improvement of the cortical map in primary visual cortex by processes analogous to spatial and temporal interpola-tion; the detection of linking features in this map; and the concentration of this information by non-topographical mapping in adjacent visual areas.

A hundred years have passed since Sir David Ferrier did his work on the localization of function in the cerebral cortex, and fifty since these lectures were founded in his honour, but even so our understanding of the overall action of the cortex has been less touched by science than anything else that influences our daily life to the same extent. I think this must be true, for if you wish to predict the next cortically controlled action that I, or anyone else, is going to make you will gain little help by consulting the scientific literature, and the youngest and most ignorant is likely to guess as well as the oldest and wisest.

In this lecture I shall be considering how the cortex deals with sensory information, rather than how it controls motor actions. As we look out through our eyes our brains, almost miraculously, give us reliable knowledge of what goes on in the world around us. I am going to attempt the reverse: to look in through the eyes at the working of the cerebral cortex, which achieves the near-miracle. This is certainly an over-ambitious attempt, for the cortex is a horrifyingly complex organ, and it is not at all a well understood task to gain reliable knowledge from messages of the type that the eye supplies. But the attempt seems worth making because we know a good deal about the physics, physiology and anatomy of some of the early steps in the process, and so we may be able to clarify the task that the visual cortex must perform.

The theme of this lecture is that, to understand the biological mechanisms of vision, we need to know the natural difficulties and limitations of the tasks being performed at each stage. Thus there are physical limits to image formation and transduction and I shall try to show how important these are in determining the structure and function of eyes. Limiting requirements are also known for the communication of the information in the image to the brain, and we are beginning to see how these limits influence the structure and function of the visual pathways. It must be equally important to identify the difficulties in the representation and interpretation of visual information in the cortex, and we shall only understand how vision gives us reliable knowledge of the world when we can identify these limiting factors with certainty and see how the biological mechanisms are adapted to deal with them.

LIMITING FACTORS IN THE DESIGN
OF THE PERIPHERAL VISUAL SYSTEM

First consider an animal whose habits cause optical difficulties. *Anableps anableps* is a fish, shown in figure 1, that lives right on the surface of the water. It presumably feeds on prey that inhabit the surface layers, and the problem of forming an image for both media is not easy because the refractive power of the cornea is abolished in water. Nonetheless, the problem is solved, as shown in the right half of figure 1.

This is only one example from the enormous variety of ocular adaptations to different visual environments documented by Walls (1942) in vertebrates. Our respect for the range of optical instruments occurring in nature is greatly extended

by including invertebrates, where one finds simple eyes in which the image is formed by reflexion in a concave mirror (Land 1965), as well as much evidence of good design in compound eyes. For instance, when the overall size of the eye changes, the size of each individual ommatidium is apparently scaled in accordance with design principles dictated by diffraction (Barlow 1952; Kirschfeld 1979; Horridge 1980). Another example of ingenious design is the superimposition of
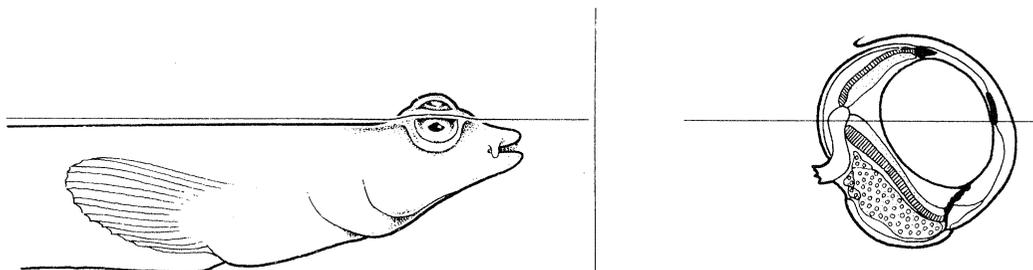

FIGURE 1. The four-eyed fish, *Anableps anableps*, and its eye. An example, from the monograph by Walls (1942), to illustrate evolutionary adaptation to a difficult optical environment.

images from individual lens cylinders according to Exner's principle (Horridge 1975; Land 1980). A final example is the recently discovered superposition eye using quite a new principle, namely an array of reflecting prisms (Vogt 1975; Land 1976).

Visual proficiency has enormous survival value and these examples leave no room for doubts about the versatility of naturally occurring optical designs. One can gain an impression of how good these designs are in detail by seeing how closely they approach the absolute limit to resolution set by diffraction and the absolute limit to sensitivity set by the quantum nature of light.

Resolving power

The right half of figure 2 shows the shape of the vitreal surface of the retina at the foveal pit of a hawk. Light from the lens may be considered almost parallel since it converges very little over the distance of a few hundred micrometres shown here. The retina, however, has a higher refractive index than the vitreous, and thus the light is diverged at the foveal pit. Snyder & Miller (1978) calculated that the spherical surface at the tip of the pit would cause a magnification of × 1.5 by the level of the oil droplets of the cones. But the fovea does better than a simple spherical lens: the left half of the figure shows an aspheric surface that gives a uniform magnification, computed using the refractive indices and distances indicated. It will be seen that the actual foveal pit conforms well with this ideal shape up to a diameter of about 60 μm; as suggested by Walls (1940), the steep, almost conical, walls of the pit also expand the image.

The hawk's acuity is 2–2½ times better than man's (Shlaer 1972; Fox *et al.*1976), and it is interesting to speculate on the ultimate factors limiting the evolution of

high acuity. It is easy to verify that the highest resolving power achieved by man (nearly 60 cycle/deg) is quite close to the limiting resolution (D/λ cycle/radian) set by the diameter of his pupil (D, about 2.5 mm in bright light) and the wavelength of light used ($\lambda = 560$ nm). But this is only a beginning: why should not the pupil be larger? One possibility which I used to favour (Barlow 1964) is that the choice
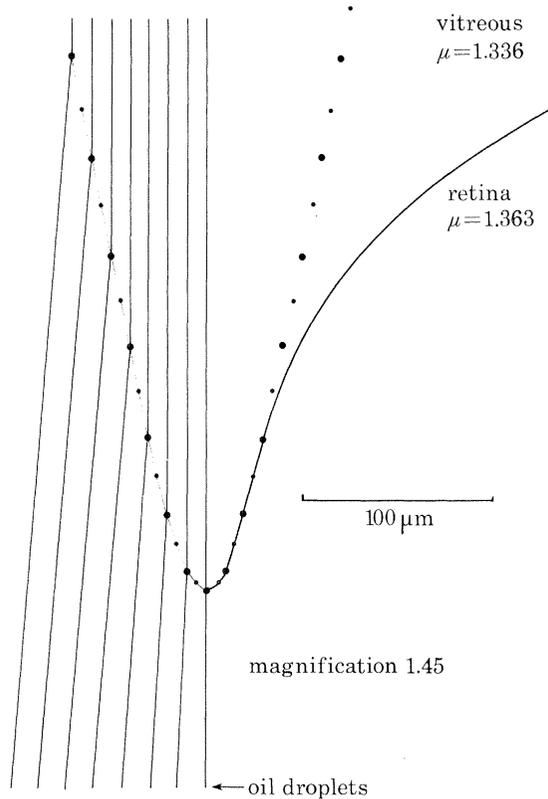
FIGURE 2. The optical principle of a hawk's foveal pit. The right half shows a tracing of the surface separating the vitreous and the retina, which has a slightly higher refractive index (μ). The dots and the thin curve in the left half show an aspheric surface computed to diverge parallel rays to produce a magnification of × 1.45 at the level of the cone oil droplets lying 100 μm beneath the tip of the foveal pit. The retinal surface will produce uniform magnification over a circle about 60 μm in diameter; the central 20 μm results from the nearly spherical surface at the tip, as shown by Snyder & Miller (1977), the remainder from the almost conical walls, as suggested by Walls (1940).

represents a compromise between a larger diameter, where chromatic aberration would be dominant, and a smaller diameter, where this and other aberrations would be reduced but diffraction at the pupil would become limiting. On the other hand, Snyder & Miller (1977) believe that limiting acuity is set, not by image quality, but by the optical 'cross-talk' between receptors that would occur if they were smaller and packed more closely together than about 2 μm. Light travels

down the outer segment of a receptor as it does down a waveguide, and since there is only a limited refractive index difference some of the energy is available in the medium surrounding the outer segment and would leak to the neighbouring receptor if this lay too close to it (Snyder 1972). On this view limiting retinal resolution, expressed in linear units, is the same in all eyes and results from the
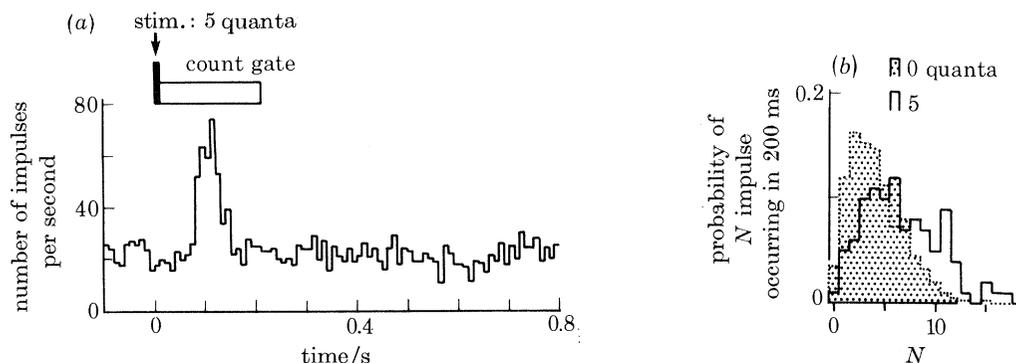

FIGURE 3. Sensitivity of the cat's eye. Flashes (stim.) were delivered containing five quanta on average, and spikes were collected from a single retinal ganglion cell to form the histogram of 100 responses shown in (*a*). The histograms in (*b*) show the numbers of occasions on which N impulses occurred in 200 ms immediately following a stimulus, and also during the maintained discharge. The average number was 4.14 ± 2.7 during the maintained discharge, and was raised to 6.62 by the stimulus. Of the 5 quanta at the cornea only 1.25 would be effectively absorbed; so on average two impulses result from absorption of a single quantum. (From Barlow *et al.* 1971.)

properties of light passing through a set of waveguides. As expected on this view, cones smaller than 2 μm have never been found, even in very small eyes, and Snyder & Miller found remarkably constant cone diameters in a series of birds of prey covering a large range of sizes. They also found pupil diameter to be directly proportional to focal length in this series, whereas the hypothesis that chromatic aberration and diffraction form the limit would predict that pupil diameter should increase with (focal length)$^{\frac{1}{2}}$. It thus seems probable that many of the structural details of an eye are dictated by the need for the outer segments of the receptors to be optically isolated from each other to resolve the image.

Sensitivity

The high sensitivity of the eye has been appreciated for many years, first from work on the human threshold (Hecht *et al.* 1942), then by direct recording from retinal ganglion cells in the cat (Barlow *et al.* 1971), and in the last year or two by measurements on single receptor cells (Baylor *et al.* 1979). Figure 3 shows the response of a cat retinal ganglion cell when flashes of light containing on average five photons were sent into the eye; as can be seen, these caused an average of $2\frac{1}{2}$ extra impulses. The fraction of photons entering the pupil that cause photochemical change in the receptors is unlikely to be more than 25 %; hence each

effective photon probably caused at least two impulses from this ganglion cell. Statistical analyses of the spike trains confirmed this conclusion (Barlow *et al.* 1971).

Figure 4 shows that single photoreceptors do indeed respond to individual photoisomerizations. It is a record made by Baylor *et al.* (1979) of the responses of a toad rod to flashes of light that caused an average of just under one photo-
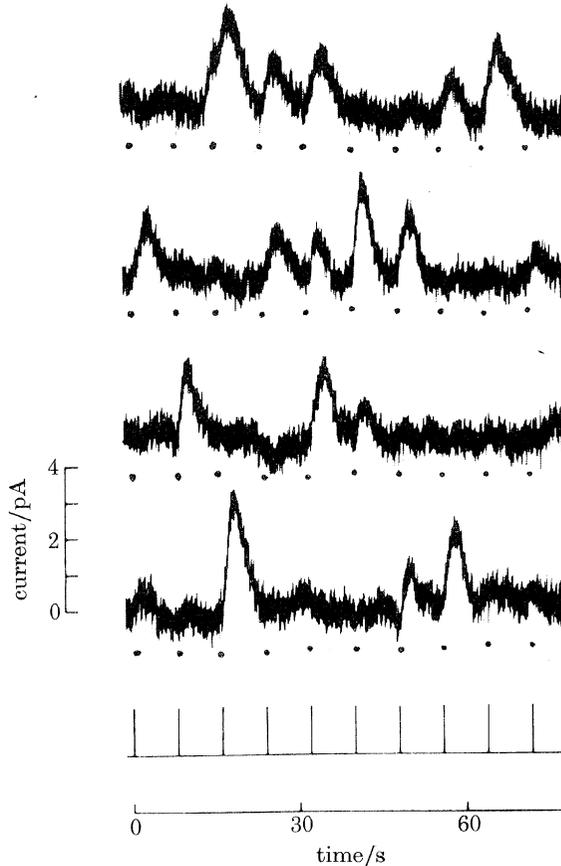
FIGURE 4. Responses of a toad rod to single quantal absorptions. Dots indicate the times of delivering a flash of light causing an average of 0.92 isomerizations of the photosensitive pigment. The bumps that follow result from decreases in the current entering the rod outer segment. The variability in the size of the responses follows the expected Poisson statistics. (From Baylor *et al.* 1979.)

isomerization. The bumps represent transient reductions in the current entering the receptor outer segment; the responses are variable because the flash caused sometimes none, sometimes one, sometimes two or more, photoisomerizations, and again, the frequencies followed the expected Poisson statistics.

It might be thought that the limit of sensitivity would be determined simply by the density of photosensitive pigment in the photoreceptors and hence the probability of photon capture, but another possibility is that the intrinsic noise of

the retina sets the limit, and that this results partly from thermal isomerizations of the photosensitive pigment (Barlow 1956). This conjecture was consistent with the psychophysical evidence, and it has recently been shown that bumps like those of figure 4 occur in the total absence of light (Yau *et al.* 1979). For the brain to be sure that light has fallen on a patch of retina it must be necessary to determine the total number of bumps occurring in many rods over a certain time, and then to see if this number is elevated significantly over the background rate. The slow rate of occurrence in the dark corresponds to a half life of the photopigment of about 1000 years at 20 °C (Baylor *et al.* 1980), which is consistent with data from humans (Barlow 1956).

It is interesting to consider thermal isomerization in cones, for they behave psychophysically as if they have a high level of intrinsic noise or 'dark light' (Barlow 1964), and evidence from electrical recording also supports the idea that they are much noisier than rods (see Barlow & Fatt, pp. 301–304 and 351–353). Furthermore one would expect reduced activation energies for photoisomerization to be associated with the shifted spectral sensitivities of red and green cones, and this would probably lead to thermal isomerization rates several thousand times higher than for rods (Barlow 1957). A high noise level would also make the morphological differences between rods and cones intelligible for in cones it would be advantageous to concentrate light optically into as small a space as possible, thereby reducing the number of noise-generating molecules while maintaining photosensitivity. The blue-sensitive cones should be an exception because their activation energy should be higher even than that of rods (Barlow 1958), and one therefore expects them to have a very low noise level; it will be interesting to learn if their electrophysiology confirms this.

This brief discussion has shown that the ability to confine light within wave-guide-like outer segments, and the thermal stability of the photosensitive pigments, are probably the critical factors limiting the evolution of better resolving power and sensitivity. Once these physical limits are understood many otherwise mysterious features of the eye become intelligible, in the sense that they are seen to be the advantageous result of evolutionary adaptation. The more central parts of the visual system have little that is not at the moment mysterious, and so it becomes especially important to seek out their principles of operation and the critical limiting factors.

TRANSMISSION OF THE VISUAL IMAGE TO THE BRAIN

Electrical engineering and communication theory may help us to understand the principles for transmitting the information in the image to the brain, because similar problems are encountered and solved in man-made instruments. It is convenient to discuss first the number of degrees of freedom, or sample points, in the image, and then go on to the requirements of dynamic range.

Sampling theory

This enables one to state how many points in the image must be measured and transmitted. Because the intensities at two sample points within one cycle of the highest spatial frequency in each dimension preserve *all* the information, knowledge of the highest frequency determines the number of points required.

The applicability of this principle receives some support from the spacing of the cones in the central fovea of man, for they are placed at approximately half the period of the highest spatial frequency in the retinal image, but complications arise when one looks at the arrangements in the peripheral retina of man or the area centralis of the cat. The image quality in the cat (Robson & Enroth-Cugell 1978) is better than its resolution (Blake 1978), and the same is true in the human periphery. Furthermore the image could not possibly be transmitted on a point-by-point basis, for there are not enough ganglion cells to sample at two per cycle of the highest spatial frequency in the image. One also realizes that a ganglion cell does not pick up from a point on the image, but from an area corresponding to its receptive field. However, both of these facts can be reconciled with good engineering principles by supposing that the resolution is deliberately degraded before transmission, perhaps partly to economize in the central processing requirements, but also with the object of improving sensitivity. The gaussian-shaped receptive field centre is then interpreted both as a means to collect more light and as a device for eliminating the high spatial frequencies, which would constitute unwanted noise. Analysis of the responses of retinal ganglion cells to sinusoidal gratings shows that they do achieve this (Enroth-Cugell & Robson 1966). Peichl & Wässle (1979) have shown the very elegant arrangement in the cat, whereby receptive field diameter and ganglion cell separation both increase towards the periphery, so that sampling is always done at about the interval appropriate to the highest spatial frequency passed by the enlarged receptive field centre.

Economy of central processing may not seem a sufficient explanation for the deliberate reduction of peripheral resolution that this arrangement brings about, even though this loss is mitigated by eye movements. Zeeman (1962) has suggested an additional advantage of avoiding good resolution in the periphery; high acuity would complicate the interpretation of the clues to depth from binocular disparity. If uniform resolution was maintained over a large area, the same system of interpretation could not be used both when the eyes were converged on a close object and when they were almost parallel for distant vision.

Thus sampling theory does give some insight into the principles of image transmission, but there are aspects of retinal physiology that do not fit in well. For instance Wässle *et al.* (1981) have shown that the cat's area centralis is catered for, not by a single type of ganglion cell spaced at the appropriate separation for highest resolution, but by both on-centre and off-centre cells of the X type that subserve high resolution, each apparently sampling at an interval that would allow the highest resolution of which the cat is capable. Nor is this all: there are

other, Y-type, cells which have poorer spatial resolution because they have larger receptive fields, and these also cover the area centralis but at the larger spacing appropriate to their own resolving power. There are also still other types of ganglion cell subserving other functions, such as signalling the direction of movement of the image over the retina, and these pick up from the same receptors in parallel with the X and Y types. Perhaps these arrangements should be taken as a hint from Nature that there is more to the problem of transmitting the image to the brain than can be understood from sampling theory. Let us therefore go on to consider a complementary aspect, the dynamic range of the information at the sample points in the image, and the available dynamic range of the nerve fibres that have to transmit it.

Dynamic range

The eye as a whole works over a vast range of luminances, from roughly 10^{-7} cd/m^2 for the lowest visible extended surface to 10^4 cd/m^2 for the brightest patches that one commonly encounters; of course much higher luminances are met if one considers the emitting surfaces of light sources such as tungsten filaments or the Sun. If one allows for changes in pupil diameter and for optical losses one finds that 10^{-7} cd/m^2 corresponds to an average absorption of about 5×10^{-3} photon s^{-1} for one rod and 10^4 cd/m^2 corresponds to about 2×10^5 photons s^{-1} for one cone. The central cones each need a separate transmitting link, which presumably must respond in $\frac{1}{20}$ s or so to avoid prolonging reaction times: hence a number of absorptions up to about 10^4 must be signalled within this time. This number is Poisson-distributed, so that its square root will be approximately normally distributed with a standard deviation of one-half at all mean rates, and it will be seen that the number of levels separated from each other by one standard deviation is of the order of 200. If one considers rods as well as cones and takes the largest receptive fields, covering several square degrees, the number of levels required is much more, but it is difficult to give even an approximate figure because of complicating factors such as the saturation of the rods, the reduced population of cones in the periphery, and the effects of a spatial weighting function representing the receptive field. However it turns out that 200 distinguishable levels are almost certainly more than a single nerve fibre can transmit in $\frac{1}{20}$ s, though it would be a very modest number for a man-made communication link. A retinal ganglion cell communicates with the brain by varying the number of impulses that it transmits in some interval of time, and the maximum number in $\frac{1}{20}$ s is no more than about 40 (Kuffler 1953); it is therefore difficult to see how the 200 distinguishable messages could be achieved. The mismatch would be worse for the peripheral receptive fields.

It seems then, that the narrow dynamic range of nerve fibres poses a serious problem, for each part of the image that must be signalled potentially contains more information than a single fibre can carry to the brain. However, we have not yet looked at the way in which the information is coded by the retina, and it turns

out that we can interpret these codes as means of overcoming the problem. For
instance the familiar arrangement of 'on' and 'off' fibres suggests that each piece
of the image is served by two fibres, one for use when the amount of light is above
the mean value, one when it is below; clearly this doubles the available dynamic
range and fits in with the findings by Wässle *et al.* of a double coverage of the cat's
area centralis by on and off X-type ganglion cells.

The narrow dynamic range of optic nerve fibres may also explain the advantages
of other features of retinal transduction. It has long been recognized that sensitivity
is automatically changed with the adaptation level in the manner suggested by
Craik (1938), and this can be regarded as a necessary adjustment to make the
levels of illumination in the image fit into the restricted range of signal levels
available in the optic nerve fibres. Similarly lateral inhibition, and also colour
opponency, can be thought of as arrangements that reduce the range of signals to
be transmitted without reducing sensitivity to local spatial or spectral differences
of illumination. Signalling of differences has this beneficial practical effect only
because there are spatial and spectral correlations in the images that are to be
transmitted: the activities in two neighbouring receptors at the same moment are
more likely to be similar than are the activities in two distant receptors at different
moments. However, we are now concerned with the fact that differential signals
have a reduced dynamic range, rather than with the reason why this is so.

The narrow dynamic range of nerve fibres as communication links seems likely
to be a critical factor in the functioning of the nervous system. Although such a
limit is less basic than those resulting from the physics of image formation and
light absorption, I think that it helps us to interpret the arrangements that we
find in the visual system, and thereby improves our understanding of it. The point
may be brought out by comparing this type of 'explanation' of lateral inhibition
with that sometimes advanced.

If critical limiting factors are emphasized one says that lateral inhibition, colour
opponency, and the gain changes of light and dark adaptation, are necessary to
transmit information about the light intensities in subdivisions of the visual image,
because the available information has a much wider dynamic range than can be
transmitted directly down a nerve fibre in a reasonable time interval. A common
alternative approach tries to explain the functional arrangements in the retina in
terms of perceptual phenomena. For instance we often use contours to recognize
objects, and lateral inhibition in the retina tends to promote the activity of
ganglion cells at contours, where luminance changes rapidly. It is then suggested
that the psychological importance of contours 'explains' the existence of lateral
inhibition, but of course this puts the relation completely backwards: the brain
can only use what information the retina supplies, so that the psychological im-
portance of contours might *result* from lateral inhibition, but could never *explain*
it. Calling it an explanation distracts attention from the critical difficulties that
must be overcome to extract useful knowledge from visual images, and it is
understanding these limits that gives real insight into the organization of the
visual system.

FIGURE 5. Sections through the skulls of hawk and man. The eyes are about the same size, and the hawk's visual performance is superior, but birds lack the vastly expanded neural system for dealing with visual information possessed by man.

MECHANISMS FOR THE INTERPRETATION OF THE VISUAL IMAGE

Some gross anatomy may show how much of the visual system remains to be considered. Figure 5 shows cross sections through the heads of a bird and man. Now, for tasks limited by the difficulties that we have so far discussed, a bird with
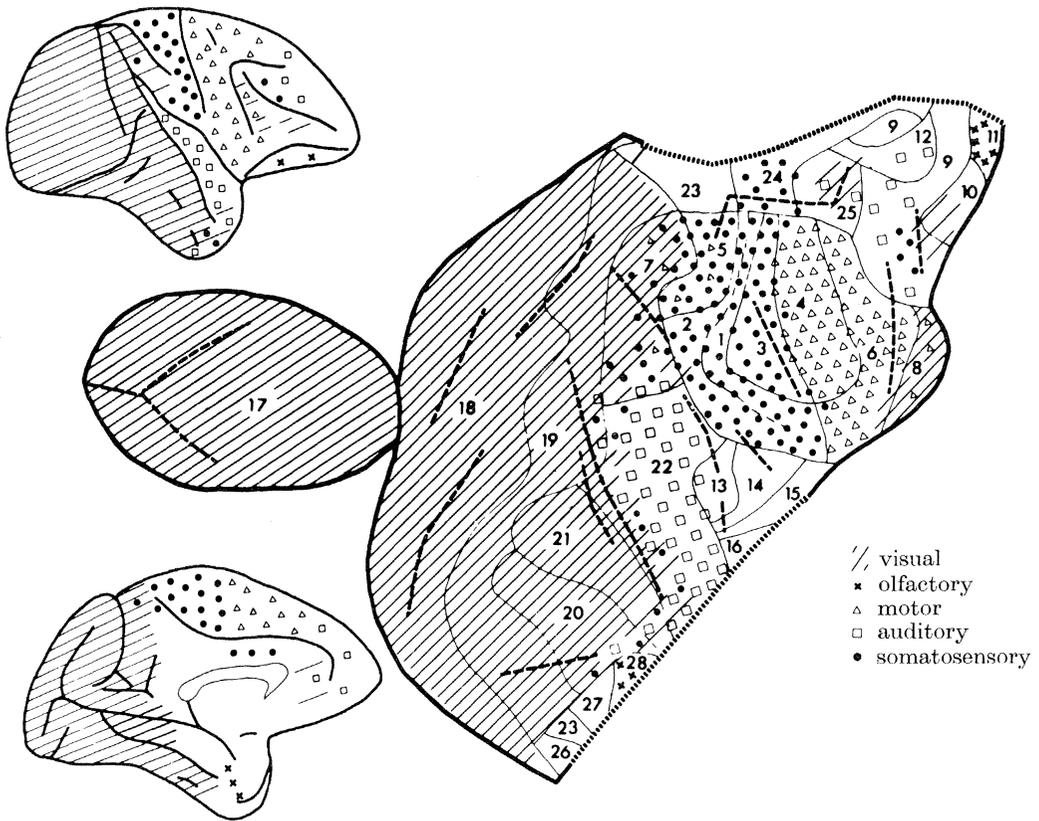
FIGURE 6. The cortex of the monkey (macaque) unfolded by Van Essen & Maunsell (1980).
The diagonally cross-hatched part is thought to be primarily concerned with vision, and
it forms as much as 60 % of the total area.

a human-sized eye is as good or better than us. But the extra space behind the eye
in the skull on the right houses much visual machinery, and we have not yet
started to discuss the problems that may arise here. Even judging by so crude a
measure as the space devoted to the mechanisms, we have not got very far in
considering the whole visual system of a primate.

Of course a monkey uses its brain for directing all aspects of its complex be-
haviour, but it is astonishing how much is directly concerned with vision. It is not
easy to judge this from a view of the visual areas plotted on the convoluted surface
of the brain, because one never knows how much lies in the sulci and hidden con-
volutions. Daniel & Whitteridge (1961) succeeded in showing the impressive size
and interesting shape of the primary visual cortex in a macaque by moulding it in
latex rubber and then pulling out the part that is normally tucked out of sight.
Van Essen & Maunsell (1980) have now devised a method for unfolding and ex-
panding onto a flat map the whole surface of the cerebral hemispheres, and figure 6
shows the result that they obtain for the macaque. Cuts and discontinuities, and
also some distortion, are required to enable the curved surface to be represented

on a plane, but they believe that it preserves the areas of the different regions with tolerable accuracy. The oval lobe at the left is the primary visual cortex, area 17, and it forms 17 % of the area of a whole hemisphere. In reality it is adjacent to the next region, area 18, along its perimeter, and this cut is the major discontinuity that they had to introduce in generating the map. In front of 18 lies 19 and these three visual areas account for nearly 50 % of the whole hemisphere, which is raised to 60 % when other predominantly visual areas in frontal cortex are included.
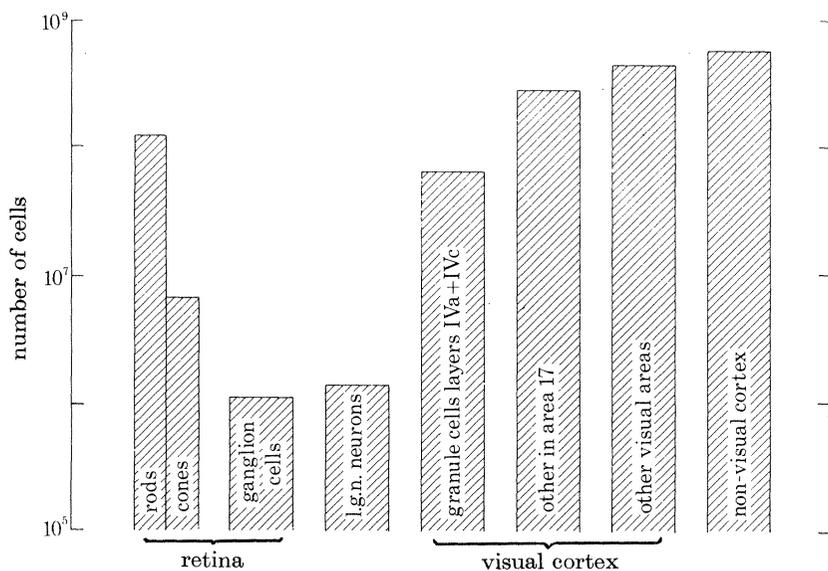
FIGURE 7. Estimated numbers of cells at various levels in the visual pathway of a monkey (macaque). Note that the scale is logarithmic, and that there are nearly 100 times as many granule cells as retinal ganglion cells. There are even more cells in other layers of the primary visual cortex, and in other visual areas. The last column on the right shows the estimated number of cells in the non-visual areas of the cortex. (Mainly from Chow *et al.* 1950.)

Some of the assignments of modality specificity are provisional, but we are forced to regard an astonishingly large fraction of the monkey's cortex as part of the visual system. Van Essen & Maunsell obtained 30 % for a similar provisional estimate for the cat, and it will be extraordinarily interesting to learn what it is for man, in whom the language faculties must to some extent reduce the dominance of vision.

These figures are intended to show that the reasonably secure understanding that we have at the level of the optic nerve takes us only a very short distance towards understanding vision. Figure 7 is an attempt to show the same point by plotting out the numbers of each class of cell in the chain leading inwards from the eye. This information is not recent (Chow *et al.* 1950), but the story told in figure 7 is very clear: the retinal ganglion cells pick up information about the

visual image from a considerably larger number of receptor cells, and they pass it through the optic nerve to a vastly greater number of elements in the cortex.

Notice that the numbers in this figure are plotted on a logarithmic ordinate scale because of the very large range involved, but when the components of a communication channel can be used in unrestricted combinations the information capacity, as usually defined in bits or bits per second, is directly proportional to their number, and therefore it would be more appropriate to expand it to a linear scale. On such a scale the numbers of ganglion cells and lateral geniculate nucleus neurons would be almost too small to show.

Information capacity also depends on dynamic range, and the figure lacks any indication of this for any of the components. It might be thought that this could change the picture, for anyone who has recorded from cortical neurons would agree that they have an unusually limited range of firing rates. However, the information capacity is only proportional to the logarithm of the dynamic range, and although one's view of the vast expansion of numbers in the cortex might be slightly modified the overall picture could not be changed to any important extent.

What is done with the vast number of cells? At the start of this lecture the visual system's task was described as that of gaining knowledge about the environment, and the first point to make is how vastly different such *knowledge* is from the mere representation that is carried by the optic nerve impulses. For instance when I say 'I see a chair' I mean that I have secure knowledge that there is a chair just beyond the reach of my left hand; if I want I can lean forward and touch it or walk around and sit on it, but a mere representation does not enable you to do this sort of thing. A lot has been achieved by the eye, the retina, and the optic nerve in getting a good representation of the retinal image into the brain, but it is the conversion of a representation into real knowledge that constitutes the barrier to our understanding.

Now when trying to think how the cortical representation of the visual image might yield knowledge of the world around us there is a very strong temptation to employ the concept of an inner eye; this mind's eye presumably belongs to some homunculus in our brain who inspects the representation and tells us about it, saying things like 'the activity in that part of the representation corresponds to a chair which you can sit on if you please'. Of course this gets us nowhere, as Descartes recognized, and most scientists want to banish the homunculus. However, he has such a tenacious grip on our ability to conceptualize what is going on that I think it will be better to ask precisely what he is doing before declaring that he is redundant. By doing this one finds (Barlow 1979; Crick *et al.* 1981) that his first job probably has nothing to do with gaining knowledge, but is simply concerned with improving the representation.

Spatial reconstruction of the retinal image

The picture supplied by the optic nerve has many gaps. It arrives in the form of volleys of impulses along discrete channels, a version suitable for transmission

but not necessarily for other purposes. Suppose, for example, that one has to decide whether three dots lie in a straight line, or whether their relative positions in the two eyes are identical; without further processing the best estimate of each dot's position would be provided by the sample element with the highest activity, but this is far from the best that could be done, and it is an experimental fact that it can be done with much greater accuracy. Human subjects can align three dots with precision better than 5′, and can detect differences in alignment between the two eyes with even higher accuracy. These are just two examples of tasks involving the determination of the relative positions of marks or objects in the visual field that can be done with a precision much finer than the separation of the sampling points in the retina. This well known fact has often caused speculation about averaging in space or time (Anderson & Weymouth 1923), but as far as I know no-one has hitherto suggested a mechanism.

If one is faced with the problem of determining the peak of a certain experimental variable one normally tries to fit a smooth curve, thereby reconstructing values of the variable between those that have actually been measured. Where the variable has been measured at points lying close enough together, as we believe is the case for the sampling of retinal illumination by ganglion cells, this can be done with complete accuracy. The procedure is simply to erect the appropriate two-dimensional interpolation function at each sample point with an amplitude equal to that of the sample, and then summate all the overlapping values of the functions at each point in the plane.

How might nerve cells perform this task of smoothing the sample points? Figure 8 shows a section through the granule cell layers of the cortex of a monkey, which receive input from the eye (Hubel & Wiesel 1972). The magnification factor is known, and so one can calculate the distance in the cortex that corresponds to a given angle in the visual field, and with use of this conversion factor the dots at the top of the figure represent the separation of the sample points required. As we have seen these need to be spaced at about one per half cycle of the finest grating that can be resolved, and it is clear that there are a very large number of cortical cells per sample point. This of course confirms the numbers given in figure 7. To the right of figure 8 a vernier is drawn with a displacement corresponding to 5′ in the visual field, using the same conversion factor. This distance is quite large compared with the average separation of cells, and it is not hard to imagine the displacement being detectable in a representation with the fine grain size and small separation of these cells. To make such a fine-grained map all that is required is that each granule cell form graded connections with a group of input fibres, the strengths of the connections corresponding to the desired interpolation function. The situation is complicated by the interleaving of the maps from the two eyes, but it will be interesting to learn if the mapping in layer IVc is orderly and regular enough for the cells to be performing this task in the simple and direct manner that seems possible.

Temporal reconstruction of the image

The homunculus also has to face the problem that the image is often moving continuously, but is only represented by impulses at discrete moments in time. In these days he often has to deal with visual images derived from cinema screens and television sets that represent scenes sampled at quite long intervals, and we know
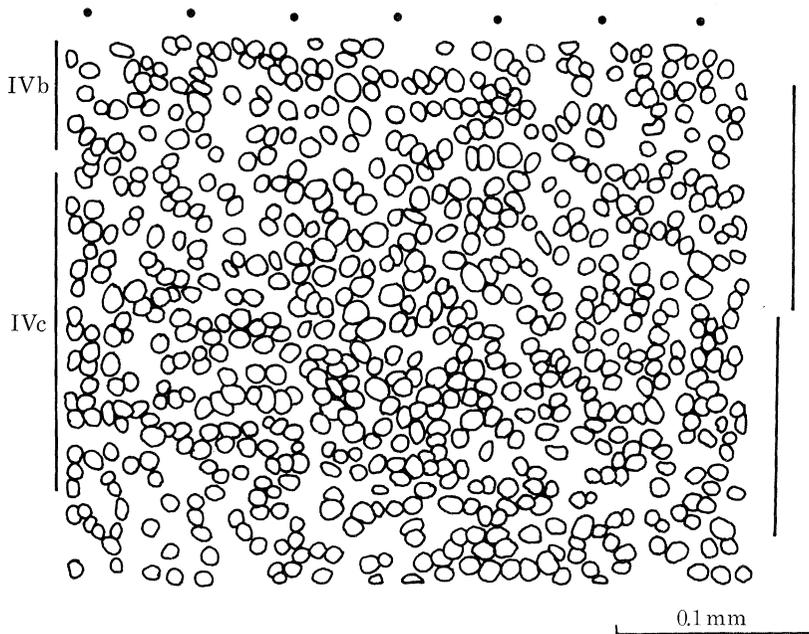
FIGURE 8. A tracing of the outlines of the granule cells of area 17 in layers IVb and IVc of monkey cortex, where the incoming geniculate fibres terminate (from fig. 3c of Hubel & Wiesel 1972). The dots at the top indicate the calculated separation of the sample points coming in from the retina, allowing two per cycle of the highest spatial frequency resolved. The misaligned vernier at right has a displacement corresponding to one-sixth of the sample separation, or 5' for 60 cycle/deg optimum acuity. The 'grain' in the cortex appears to be much finer than in the retina.

that he does a good job at interpreting them even when the sample rate is only 16 s^{-1}, as in amateur movies. One only has to watch a kitten playing, a cat hunting, or a bird alighting at dusk among the branches of a tree, to appreciate the importance and difficulty of the visual appreciation of motion. Considering this overwhelming importance it is surprising to find how slow are the receptors and how long is the latency for the message in the optic nerve, and even more surprising to find how well the system works in spite of this slowness.

Recent psychophysical work has improved our understanding of these problems. At one time it was thought that image motion aided resolution (Marshall & Talbot 1942), but this was hard to believe because of the blurring effect of the eye's long

integration time, and it was disproved eventually by measurements of acuity under conditions where the motions of the image over the retina were compensated and thus the image was held stable in one position (Tulunay-Keesey 1960). Westheimer & MacKee (1975) therefore asked the interesting alternative question 'How well does the eye *resist* the expected degradation of performance by image motion?'. For tasks such as vernier acuity they found remarkably little change provided that the rate of movement was 2 or 3 deg s^{-1}. During the integration time of, say, $\frac{1}{50}$ s, the image moves across about six of the sampling positions required for resolving the finest gratings, and a distance at least 30 times the values for positional accuracy that they actually measured. At the very least this must mean that the timing of a change in the retinal image is preserved much better than one might expect from the sluggish characteristics of the receptor responses.

Some recent experiments begin to show how this timing information is used. Everyone who has watched a movie or television knows that an impression of smooth movement is given by a succession of stationary images presented at intervals of $\frac{1}{20}$ to $\frac{1}{50}$ s. I had always supposed that the impression of smooth motion from what is in reality a jerky staircase must involve processes of temporal and spatial integration that would lead to the loss of a great deal of resolution. Suppose that one took a cine picture of a man walking from left to right using two cameras, one for his head and one for his body, and then projected them with two projectors. Obviously if the projectors were properly aligned and the films started at the right moment one would expect to get a good picture of the scene, but recent experiments have shown that this would actually be extremely difficult because of the precision of temporal synchrony that would be required: a delay of a millisecond in one projector would cause the head or body to be seen displaced relative to its companion. It looks, then, as if the exact timing is utilized to reconstruct the trajectory of an object by interpolation, for why else would a small temporal delay cause the impression of spatial misalignment?

These conclusions come from experiments done more or less simultaneously in several laboratories (Burr 1975, 1979; Morgan 1975, 1976; Ross 1974; Ross & Hogben 1975; Morgan & Thompson 1975; Burr & Ross 1979). The initial observation was that the false impression of movement in depth, known as the Pulfrich effect, occurs with intermittent illumination. The conditions for the appearance of movement in depth in the Pulfrich effect are first that the object should be moving across the visual field and second that the image in one eye should be dimmer than that in the other. It results from the fact that the latency is a little longer from the dim image than from the bright one, thus introducing a discrepancy between the eyes in the apparent position of the object at any one moment; this discrepancy is then interpreted by the stereoscopic mechanism as due to a difference in depth. With stroboscopic illumination the object must be illuminated at the same position in both eyes, but the false illusion of depth still occurs and suggests that, for a moving object, a difference in time of arrival of messages in the brain is interpreted as a difference in position. There were loopholes in this interpretation arising from

the possibility that the subject's eyes tracked the moving object, but they were eventually overcome both by direct measurement (Morgan & Turnbull 1978) and by designing the experiment in such a way that tracking movements could not start before an observation period was complete (Burr 1979).

There are two results that are worth giving in more detail because they tell us something more about the mechanism. Burr & Ross (1979) had an experimental set-up in which subjects viewed an intermittently illuminated moving object at the same brightness in both eyes, but with independent timing for each eye. To create an impression of depth they could either introduce a real spatial disparity, or a temporal delay in one eye that gave the same effect. They could also nullify a spatial disparity with a temporal delay, thus proving that time and position were completely interchangeable in giving the cue to depth. The delay required to annul a spatial disparity was exactly that predicted from the angular velocity of the object, and no clearer proof can be asked that apparent position is determined by a mechanism of interpolation.

Their results also enable the accuracy of utilization of timing information to be assessed, and it is of the order ± 150 μs; this is an astonishingly low figure considering that the latency, even at the ganglion cell level, is rarely less than 100 times as long (Kuffler 1953). A very clear demonstration of how unexpected this result is can be deduced from the fact that no neurophysiologist that I know of records the time of cortical impulses with an accuracy greater than 1 ms and they are usually grouped in bins of 10–20 ms to form histograms of responses. The brain certainly uses timing information an order of magnitude more accurate than is available from such a histogram, and so we may be missing something important.

The second result of interest is from an experiment in which Burr (1979) measured the accuracy of alignment of an intermittently illuminated moving vernier. Time and position were interchangeable, as expected from the earlier experiments, but he showed something more. To prevent tracking, the moving vernier was only displayed for 125 ms, and it could move in either direction. It was, however, only at the end of this display time that the interchangeability of time and position could be demonstrated. No eye movement could have started at this time, but it is as if some internal spatio-temporal integrating process had to be initiated by the first 100 ms of movement before a delay was interpreted as a change of position.

The mechanism for temporal interpolation must lie early in the visual pathway, and, as we have seen, it must achieve the high positional accuracy that is required, not only for detecting displacements in moving verniers, but also for such tasks as bird-catching or playing tennis. It has already been pointed out that the representation provided by the optic nerve is quite complex, and details of the termination of different classes of fibre in the granule cells of layer IV suggest that the Y or transient fibres terminate in different layers from the X or sustained fibres (LeVay & Gilbert 1976). Thus there are probably two reconstructions, one of rather poor spatial resolution that is rapidly formed from a moving image and

rapidly erased when it has passed, the other of higher spatial resolution and longer persistence. Perhaps this is a clue to the interpolation mechanism, for to take account of movement one needs to combine information from two different places at two different times. The two reconstructions might effectively apply to two different times, so that the problem would be simplified to that of detecting
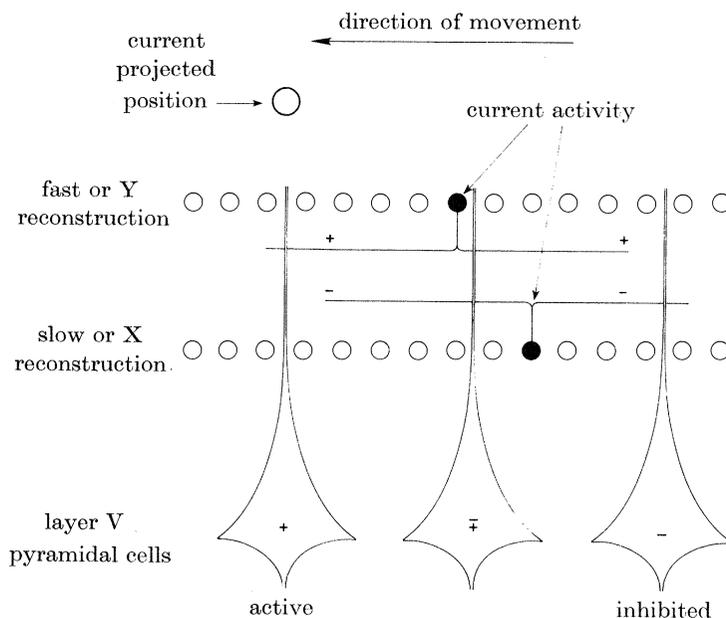
FIGURE 9. Hypothetical scheme for detecting direction of motion and predicting future position of a moving object. The more persistent representation formed by X cells prevents the faster and more transient representation from the Y cells from exciting the apical dendrites of pyramids, except on the leading edge of the spread of axons from the Y-cell granules.

synchronous activity in X and Y maps at different places. Figure 9 shows how this might be done, and how such a mechanism might actually predict the future position of a moving object. In this scheme the detection of a moving object depends upon the avoidance of inhibition rather than the detection of a conjunction, which is in line with the mechanism for motion detection suggested for rabbit retina (Barlow & Levick 1964), and for which there is also evidence in the cortex (Goodwin *et al.* 1975; Sillito *et al.* 1980). Measurements of contrast sensitivity made with stationary and rapidly moving gratings (Burr 1981) also suggest that the system's good performance for moving objects is achieved by arranging that such targets pass unscathed through a screen of spatial and temporal inhibitory mechanisms.

A developmental factor?

To set up the system for spatial and temporal interpolation envisaged above would require very precise adjustment of the synaptic potencies of the input fibres to the granule cells. We know from the developmental work of Hubel & Wiesel (1970, 1974) that, without any visual experience at all, connections are made that map the retina on to the cortex with considerable precision, and also endow cortical neurons with the rudiments of their normal, adult, pattern selectivity. But most workers in this area would now agree that cortical neurons are far from normal in their responsiveness and selectivity unless they receive a good many hours or days of visual experience during the critical period up to 3 months within which they are sensitive to deprivation. Is it possible, then, that the fine adjustments of connectivity require experience, and that spatial and temporal interpolation could not be done so precisely without these adjustments?

There does not seem to be any theoretical necessity for the fine adjustments to require experience. The fine-grained map is merely a highly redundant form of the point-sampled representation carried by the input fibres. The redundancy consists in the strong correlations between activities of neighbouring elements, and it would not necessarily require external information to set this up in the right way. One can, however, imagine a system in which external activation brings about adjustments that compensate for imprecision in a developmentally laid down pattern of connections. This would be similar to the suggestion made by Platt (1960) to account for the high precision of positional judgements and by Blakemore & Van Sluyters (1975) to account for the alignment of the preferred orientations in the two eyes of binocular cortical neurons.

A good many recently discovered facts fit in with the notion that experience has such a positive role (Imbert 1979; Rauschecker & Singer 1980). For instance Derrington & Fuchs (1981) and Bonds (1979) have shown that contrast sensitivity functions in deprived eyes tend to lack the low-frequency cuts found normally, suggesting that the development of appropriate inhibitory connections requires experience. In particular Freeman & Bradley (1980) have shown that amblyopic eyes have disproportionately poor vernier acuity, whereas the mate to an amblyopic eye often has better vernier acuity than normal; this is a striking example of the positive role of experience, and specifically implicates interpolation mechanisms. The prediction is, then, that visual experience during the sensitive period up to 3 months is required to set up the detailed connectivities that enable the properly reconstructed image to be formed. It would follow that, if this process is prevented by visual deprivation, the types of performance that would be specifically disturbed are those that require spatial and temporal interpolation. Alignment of a moving vernier should be a particularly difficult task for an amblyope to perform.

If these ideas are on the right lines the first task performed on the representation of the visual image in the cortex is the formation of a highly redundant fine-

grained map that can be used for those subsequent operations that require information about the image at positions and times between the samples provided. It must be remembered that colour has not been considered at all, that only the sketchiest ideas about the cellular mechanism have been advanced, and that even the term 'reconstruction' is a bit of an exaggeration, for the low spatial and tem-
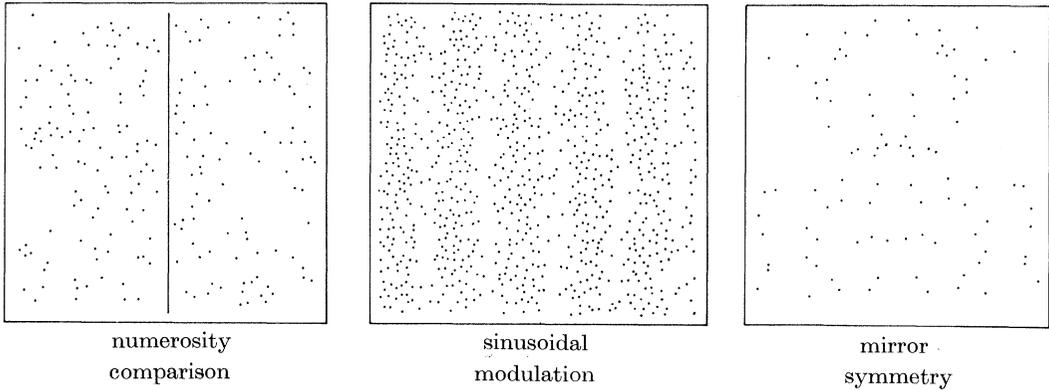

numerosity comparison	sinusoidal modulation	mirror symmetry

FIGURE 10. Subjects use up to 50 % of the available statistical information when detecting the ordered features embedded in patterns of random dots shown above (from Barlow 1978; Van Meeteren & Barlow 1981; Barlow & Reeves 1979). To achieve this level of performance efficient mechanisms for collecting the relevant information and making statistical decisions would be required.

poral frequency information in the retinal image has not been restored. But, interesting as it is, the task of reconstruction is not the one for which the inner eye of the homunculus seemed so essential.

KNOWLEDGE FROM CORTICAL MAPS

What we require the homunculus to do next is to collect together the information from scattered parts of the visual image and to tell us about the chair or horse that this information enables him to recognize. The essential step in obtaining such knowledge from the sets of data that are contained in the cortical maps is that of making a statistical decision. The visual system has to deal with fluctuations and uncertainties in the scenes that it surveys, and because of these uncertainties there are definite limits to the reliability of any decision. It is important to ask how closely the overall system approaches these limits of statistical inference because this will tell us how well the job of collecting together the information is performed.

To answer the question it is necessary to choose tasks for which the uncertainty can be accurately quantified without making any prior assumptions about the properties of the visual system that is making the inference. One way of doing this is to generate patterns, composed of random dots, containing one consistent

feature that has to be detected in spite of the uncertainty created by the random element. Figure 10 shows examples of such patterns in which subjects had to detect differences of average dot density (left), the sinusoidal modulation of density (middle), or the symmetrical disposition of dots (right). The upshot of these experiments (Barlow 1978, 1980; Barlow & Reeves 1979; Van Meeteren & Barlow 1981) has been to show that, for tasks well matched to the properties of the human system, it performs with high efficiency, equivalent to utilizing 50 % or more of the statistical information available in the samples that are presented. Now a blind man can tell you where the Sun is by the heat falling on his cheeks, but by comparison with physical instruments his performance would be appallingly bad. For many perceptual tasks, such as symmetry detection, one could not possibly say *a priori* whether performance is really good or bad on an absolute scale, and if it were bad someone might be tempted to say that symmetrical patterns reminded the subject of his mother's face, or some such nonsense. But when a task of this sort is performed well for arbitrary random patterns one must consider carefully what mechanisms are required, for this performance cannot be an accidental property of a mechanism whose main function is something quite different. In particular the result shows that much of the information available for performing a recognition task is effectively gathered together. What we must find, then, are mechanisms for collecting together the information that enables us to do the varied types of recognition of which we are capable.

LINKING FEATURES AND NON-TOPOGRAPHICAL MAPS

The Gestalt psychologists were fascinated by the problems posed by the interactions of remote parts of a visual image, and they have given us demonstrations of some of the processes at work. They did not, however, point out the importance and difficulty of the task of collecting together the information for, as an instance, detecting symmetry, nor do their demonstrations give any hints on how nerve cells may bring this about. I think that the key to understanding how this is done lies in two concepts, linking features and non-topographical maps. Linking features

FIGURE 11. Hypothetical non-topographical maps. In the map of the *orientation* of line segments at top the ordinate is the orientation of a line segment and the abscissa its distance from the fovea measured perpendicularly to its orientation; a continuous straight line is represented by a point, a circle centred on the fovea forms a vertical line, and the row of oblique segments forms another line. In the *motion* map the direction (θ) and distance (r) from the centre (O) indicate the direction and velocity of movement; in the view from a train window to the left the eye fixates a tree in the middle distance, and this and other stationary objects go to O; the window frame is moving on the retina in a direction equivalent to leftwards, and projects as indicated; an object in the foreground moves rightward, the bird flies upward and the cigarette ash drops downwards. In the *colour* map, similarly coloured parts project to the same point, and differently coloured parts to different points. As indicated for the red parts, approximate positional information could be retained for each mapped quality if there is topographical mapping on a microscale (< 1 mm) as well as the non-topographical mapping on a macroscopic scale; area 17 has such double mapping, but there the macro-map is topographical, the micro-map non-topographical.

visual
field

non-topographic
maps

orientation

distance from fovea

Orientation

direction (θ) and velocity (γ)

Motion

green

yellow
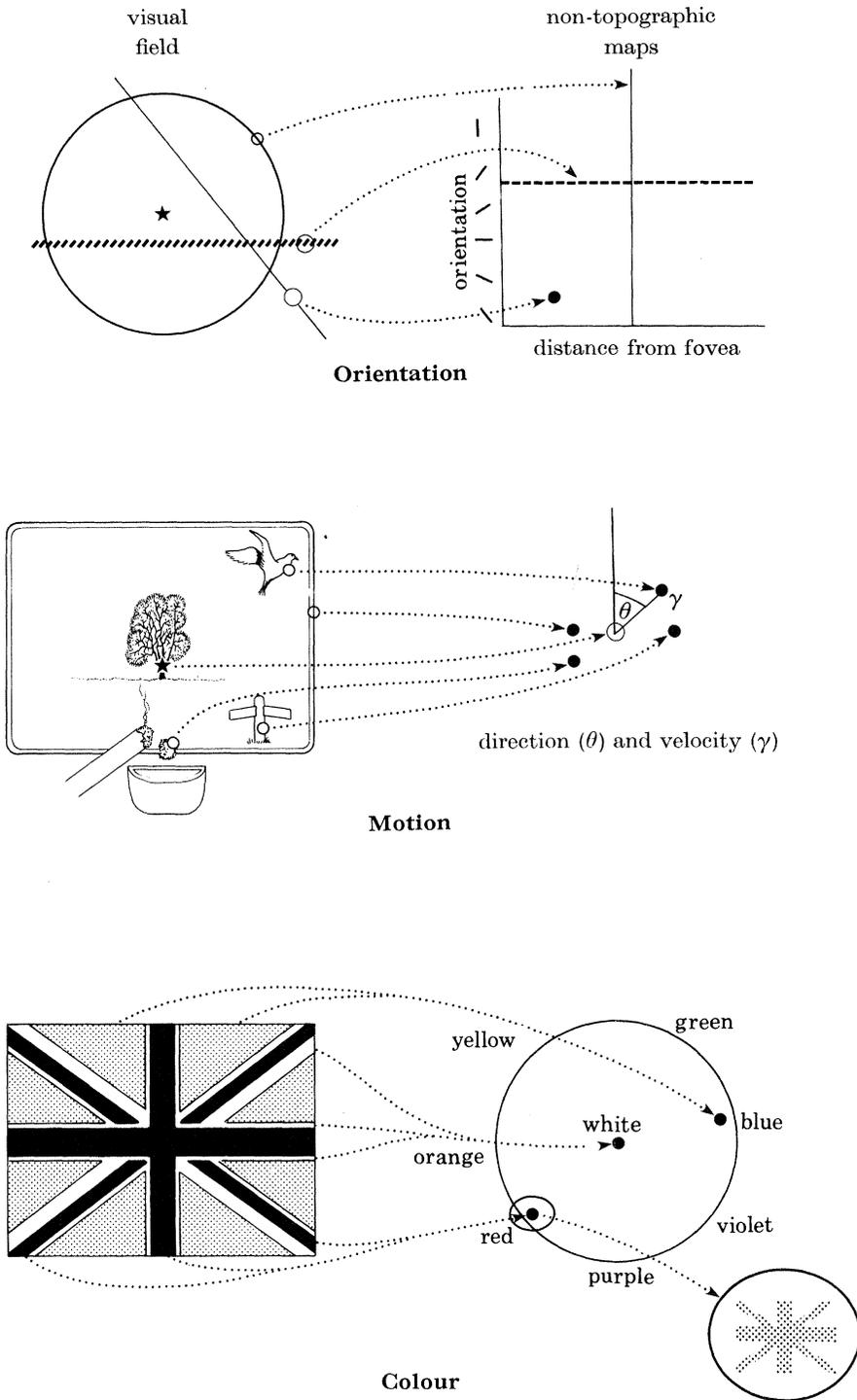
white

blue

orange

red

violet

purple

Colour

FIGURE 11. For description see opposite.

are those locally detectable qualities of a portion of a visual scene that, in Gestalt terms, cause segregation, or separation of figure from ground. Colour, texture, disparity, direction and velocity of motion, and orientation are examples. There are many quite familiar demonstrations that show what potent factors these are in causing parts of a scene to be grouped together, and their utility was especially well demonstrated in a rather different way when Guzman (1968) solved the task of programming a computer to identify and separate in a photograph several blocks of wood lying in a jumbled heap. In this case the main linking features were the edges; lines at a particular orientation and position probably belonged to the same block, and served to identify it even when partially obscured by other blocks.

Non-topographical mapping is a rather less familiar concept. In a topographical map neighbourhood relations are preserved; if two points lie near each other in the original, they will lie near each other in the map. In a non-topographical map it is not topographical propinquity that is preserved, but propinquity along some other dimension such as movement or colour, and in such a map points moving similarly or coloured similarly will lie close to each other, no matter how far apart they were topographically in the original. Figure 11 illustrates hypothetical examples of such non-topographical maps and these are explained in the legend; the need to consider them arises from recent advances in knowledge of the anatomy of the cortex.

Cortical lamination and the redistribution of the image

It is only recently, following the development of enormously improved anatomical methods for tracing the origins and destinations of nerve fibres, that the significance of the well known lamination of cell bodies in the cerebral cortex has become evident; the cells whose bodies lie in different laminae have been found to send their axons to different destinations (Gilbert & Kelly 1975). Thus in the primary visual cortex the cells of the infragranular laminae V and VI send their axons mainly to the superior colliculus and the thalamus respectively, whereas those of the granular and supragranular laminae IV, III and II send them mainly to other parts of the cerebral cortex.

Some of these projections fit current ideas about the role of pattern selectivity in visually guided behaviour. For instance the superior colliculus is thought to be responsible for the initiation of eye movements to fixate objects of likely interest in the visual field (Wurtz & Albano 1980); hence the layer V cells of the cortex may behave in a manner similar to that suggested many years ago for ganglion cells in the frog's retina (Barlow 1953; Lettvin *et al.* 1959); they would be the equivalent of 'fly detectors', but they would work from the cortically reconstructed image, not the retinal image, and they would initiate a flick of the eye rather than of the tongue. The mechanisms suggested in figure 9 might even enable some prediction to be made of the future position of a moving object so that the eye arrived at a new fixation position at the same time as the target to be foveated.

That type of analysis would certainly be interesting, but it simply involves the

detection of a specific trigger feature wherever it occurs in the visual field, and we need something more to explain how the information for global analysis of the image is collected together. We are, however, unlikely to find evidence for such global analysis in area 17 itself because of another recent anatomical clarification.

Range of intracortical connections

Until a few years ago one might reasonably have held the view that intracortical connections were so numerous and extensive that one could not confidently exclude the possibility of a direct connection between any pair of cells anywhere in the cortex; now very important restrictions are being established. Probably the most important of these for the purpose of understanding the potential function of cortex is that the very dense and numerous interconnections only extend for a millimetre or so (Fisken *et al.* 1973), and that the longer connections, though also numerous, go to specific and restricted destinations elsewhere in the brain (see also Braitenberg 1973). The dense local interconnections presumably perform the detection of local pattern features in the image, but the redistribution of this information to other areas is a second major function of area 17.

These new facts must change how we think about the visual cortex, and linking features and non-topographical maps seem to provide the keys. First one sees that it must be a mistake to try to account for the global aspects of perception in terms of what goes on in area 17 by itself. Because of the limited range of connectivity it can only handle information from a small part of the visual field at one time, and so it cannot possibly collect global information. However, such local analysis is just what is needed to detect each of the different kinds of linking characteristic. The functional role proposed for area 17, then, is first to detect linking features by local analysis of the topographically organized reconstructed image, then to redistribute this information to other regions. Furthermore the details of this redistribution gain greatly in importance. If all regions with the same linking characteristic are to be detected they must surely all project to the same small region of cortex, for only in this way can the activity of remote parts of the visual field be brought within the range of 1 mm or so where intracortical connectivity is dense. Hence it is proposed that the reprojections from area 17 create non-topographical maps in parastriate areas. Those global aspects of perception that so intrigued Gestalt psychologists are thus thought to be achieved in two stages: first specific linking features are detected by local analysis in the topographical map in area 17; then information about each of them is collected in a separate parastriate region.

There is quite good agreement between the pattern selective properties of neurons in area 17 and the characteristics of a scene that are known to act as linking features and cause perceptual segregation of subsets of the image. The first is direction and velocity of motion. A scene of dots and splodges that is quite meaningless when all are stationary will suddenly reveal a human figure when the appropriate subset of marks is caused to move in one direction. From the earliest

recordings of cortical neurons it has been appreciated that many are selective for both direction and velocity of movement (Hubel & Wiesel 1959, 1962), and on the current hypothesis their functional role is to detect and classify the linking features provided by motion.

If the same subset of marks is coloured, the figure will be equally evident. The colour-selective properties of cells in the cortex are not well understood, but it seems that colour-selective cells certainly occur (Michael 1978a, b). At this point a prediction can be made, for the postulated linking character must be the surface reflectance of the scene rather than simple spectral composition, and so one would expect colour-selective neurons in primary cortex to have some of the properties of those described by Zeki (1980) in his parastriate colour area.

Another characteristic that causes perceptual segregation is disparity, as shown in many demonstrations to be found in Julesz's book (1971). Here again cells in primary cortex are selective for disparity both in cat (Barlow et al. 1967; Nikara et al. 1968; Ferster 1981) and monkey (Poggio & Fischer 1977).

Finally there are the well known properties of orientational selectivity found in cortical neurons (Hubel & Wiesel 1962). This certainly fits in with the importance of edges as linking features in Guzman's wood-block problem, and contours are important in many tasks where the objects to be segregated are less regularly shaped than they were in that case. But it is also attractive to suppose that neurons selective for size and orientation provide a general description of the texture and predominant orientation of each small section of the image by performing something akin to a local spatial frequency analysis. This is the scheme proposed by Robson (1975, 1980) and it has now been well established that many cells in the cortex have a bandpass spatial frequency response characteristic, the width of the response band being sometimes as narrow as one-half of an octave. This would not be nearly narrow enough for a Fourier transform of the whole image that contained more than a small fraction of the information in it, but a recoding of this sort could form an economical and complete way of representing all the information contained in small regions of the image, perhaps corresponding to one hypercolumn (Sakitt & Barlow 1977, 1981). Texture and orientation are of course effective means of causing portions of an image to become perceptually segregated.

Parastriate visual areas

We have seen that the pattern-selective properties of neurons in area 17 conform quite well with those required to pick out linking features. How well does the neurophysiological evidence support the idea that this information is redistributed to form non-topographical maps? From the work of Allman & Kaas (1974, 1976) on the owl monkey, Zeki (1974, 1978) on the macaque, and Tusa & Palmer (1980) on the cat, it is clear that there are many accessory visual areas that receive their input largely from area 17 and thus might receive such maps. In most if not all cases some degree of topographical mapping has been found in these areas, but all

would agree that these maps are much less precise than in area 17. One badly needs a measure of this precision, such as the number of meaningfully different loci in the region; this might be obtained by the ratio of the total area of the region to the scatter at one locus, this scatter being expressed as a cortical area by using the appropriate magnification factor to convert the scatter of receptive field positions at one locus into the equivalent cortical area. My impression is that the number of meaningfully different loci would be small in many instances, but no such measures have been used. Most of the experimentalists, who do the laborious part of the work, have so far concentrated their efforts on topographical maps and their success in showing some consistent mapping probably means that schemes such as those shown in figure 10 cannot be more than idealizations. However there is also a little positive evidence for non-topographical mapping.

Zeki (1974, 1978) stresses the crudeness of the topographical map in the accessory visual areas in the macaque. He has also shown that the different areas predominantly collect different types of pattern-selective information. Thus one region collects motion information, while another has a high proportion of colour-specific units. There is, furthermore, a hint of non-topographical mapping within each area for he shows that cells with similar properties are clustered together. This receives some support from the anatomical work of Montero (1980). In the colour area Zeki (1980) shows that it is the surface reflectance of objects that determines the response of cells, but on the view developed here this is likely to be derived from area 17, and it would be the segregation of similarly coloured regions that occurs in Zeki's area.

In area 18 of monkeys Hubel & Wiesel (1970) described cells selective for binocular disparity, and these cells were, furthermore, grouped according to disparity. This is what one would expect if depth or disparity was a dimension employed for non-topographic mapping, but the precision of the topographic map in 18 appears to be high, and so one cannot seriously propose that depth is the main mapping principle. It is worth recalling that different forms of mapping can exist in one cortical region, as shown by Hubel & Wiesel (1977) in area 17. There the coarse macromapping, accurate to about 1 mm, is topographic: by moving 1 mm or so on the cortex one moves to a distinct non-overlapping area of the visual field, but within each 1 mm there is a microstructure in that ocular dominance is arranged in an orderly way. Furthermore outside layer IV there is an orderly arrangement of the orientational selectivity of the cells. In the accessory visual areas neither macro nor micromaps are properly understood yet. Possibly their roles are reversed, colour or direction of movement being segregated on a coarse scale while approximate position in the visual field is still mapped, but repeated many times on a microscale, as Hubel & Wiesel found for orientation in area 17.

If the detection of global properties of the visual scene is accomplished by the non-topographical mapping of the information obtained by detecting local linking features in the reconstructed visual image in area 17, then one expects a separate area for each linking feature. The scheme is illustrated in figure 12.

H. B. Barlow

It will of course be asked 'What is done with the information that there is a region of the visual field with some common direction of motion, or that collinear orientation detectors are being activated?' I like the idea that this information is signalled back to the reconstruction in 17, enabling the area that has the common characteristic to be 'flashed' or 'cross hatched' in some way, though I am not sure that this idea would be so appealing if we had successfully banished the

Area 17

(1) reconstruction of highly redundant, fine-grained retinal image with centre emphasis and reduced low-spatial frequencies

(2) local detection of linking features and distribution to other parts of brain

(3) probable normalizing function

Pre-striate areas

deal with specific linking features to identify regions containing common ones

Possible linking features

orientation

spatial frequency

disparity

movement

colour

area 17

pre-striate

annectant gyrus
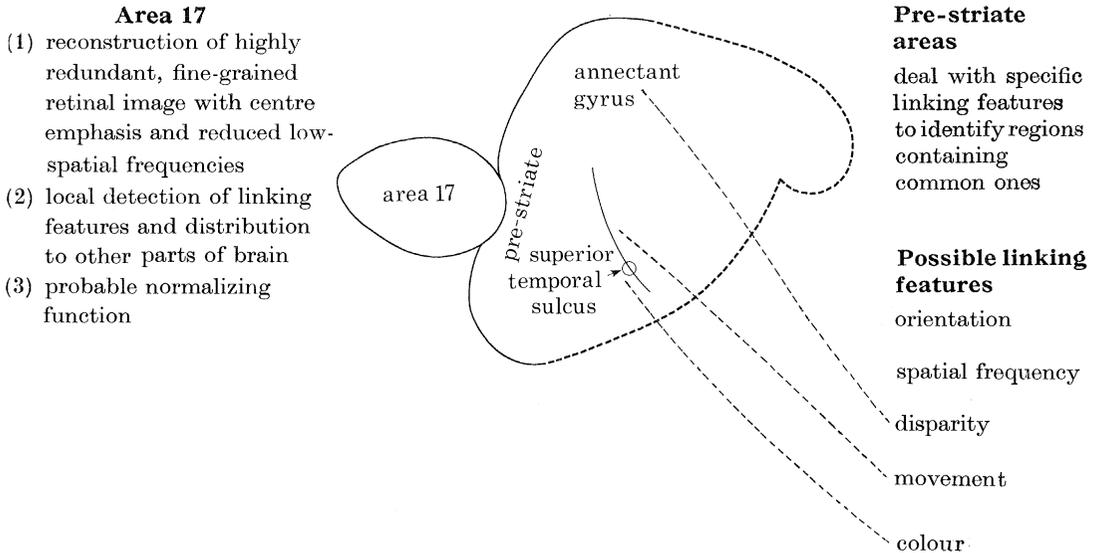
superior temporal sulcus

FIGURE 12. Suggested functions of visual cortex. The primary cortex (area 17) reconstructs the image, smoothing out discontinuities in space and time, and detects local linking features (orientation, movement, texture, disparity, colour), which are then distributed to separate regions in parastriate cortex, where non-topographical maps may be formed. The approximate location of these regions is shown on the unfolded map of Van Essen & Maunsell (1980) (see also figure 6). Evidence for reprojection of information about disparity, colour and movement has been given by Hubel & Wiesel (1970) and Zeki (1976). They also show that the representation of particular linking features is localized within these areas, but there is no evidence yet for the features being arranged in an orderly way to form non-topographical maps.

homunculus from that area. But, whatever is done with the information next, a hopeful working hypothesis of the first steps in perception has emerged from the twin concepts of linking features being detected in primary visual cortex, and perceptual segregation of the regions where they occur being accomplished in non-topographical maps in parastriate areas. Notice the critical importance of limited cortical connectivity in reaching this hypothesis.

There must be other critical limiting factors that would help us to understand the cortical mechanisms of perception. For instance, although the physiological mechanisms whereby pattern selectivity is achieved are not well understood, it appears likely that the restrictive condition, equivalent to logical conjunction, is

achieved by inhibition that is logically equivalent to 'and not' (see, for instance: Barlow & Levick 1965; Sillito 1980). This would not make any logical functions impossible to achieve but it would certainly influence the organization of the system and the ease with which different functions could be realized.

<div align="center">CONCLUSIONS</div>

I can imagine a friend summarizing my lecture in words such as these: 'First he pretended to understand the peripheral visual system; then he even pretended to understand the visual cortex.' But I do not apologize for this, because it is only by making hypotheses and then correcting them that we may eventually gain a real understanding of the visual system. To summarize the current position, critical limiting factors for resolving power and sensitivity appear to be the confinement of light within the outer segment of a single photoreceptor, and the stability of photosensitive molecules. For transmitting the image centrally the narrow dynamic range of nerve fibres appears to be most important, and in the cortex itself the difficulty of collecting together the required samples of information and the limitations on cortical connectivity are probably the critical factors.

The hypotheses that I have suggested arise from these limiting factors, and there are several predictions that should be experimentally testable. To mention some: the blue cones should be much less noisy than the red- or green-sensitive cones; the granule cells in layer IV of the cortex should have receptive fields arranged in an orderly map and separated from each other by an angular distance much less than the separation of ganglion cell receptive fields in the corresponding part of the visual field; there may be more to the representation of motion in area 17 than has yet been described, to account for the eye's remarkable ability to interpolate the position of moving objects (see figure 9); the mechanism for spatial and temporal interpolation may be especially sensitive to deprivation in the sensitive period, and thus should be specifically disrupted in amblyopia; colour-sensitive neurons in area 17 should already be selective for surface reflectance, since it is this and not spectral composition that constitutes a valuable linking feature; every character that causes perceptual 'segregation' should be detected by neurons in area 17, and projections from area 17 to parastriate regions should ultimately create non-topographical maps in which these properties are brought together from separate regions of the visual field; within each such non-topographical map position may also be mapped on a microscale in a manner analogous to the mapping of orientation and ocular dominance in area 17. Nature rarely does things the way that we expect and I fear that much of this theorizing may be wrong, but it has made testable predictions and also has an interesting general implication.

Over the last hundred years the localization of function within the cerebral hemispheres that Ferrier was among the first to establish has been a dominant idea. The demonstration of well organized cortical maps of the sensory surfaces was

not only one of the fruits of this approach but also seemed to support the notion that the role of a particular cortical region was often simply to map a particular part of a sensory surface, and thus carry a representation of the activity at that sensory surface. But this may be very misleading: a topographical map is the *only* way to represent close together in the cortex events that occur close together at the sensory surface, so that if the function of area 17 is to detect local properties of the image, a topographical map is a necessity, as Cowey (1979) has also pointed out. But, equally, for detecting global, or non-local, properties a form of representation is required that brings signals of the events to one locus in the cortex wherever they occur at the sensory surface, and to do this non-topographical mapping is required. In this way the whole cortex acquires its unity again, for it all becomes association area, and the primary projection areas with good topographical maps are simply regions specializing in the detection of local associations. This is an important change of viewpoint, for the natural question to ask about a particular cortical locus becomes 'What types of information are *brought together* here?' rather than 'What is *represented* here?'.

It is not of course a new idea that the cortex is where associations are detected and knowledge of the environment established, but it was hard to find room for these processes on the view that the cortex is simply a screen on which the events at the sensory surface are signalled. The scheme of detecting the local associations that constitute linking features, and reassociating this information in non-topographical maps may, just possibly, release us from the straight-jacket in which pure topographical maps and simple representations have tended to confine our thoughts. However we shall only know if such a release is merited after many more hard-won experimental facts are established.

REFERENCES

Allman, J. M. & Kaas, J. H. 1974 A crescent shaped cortical visual area surrounding the middle temporal area (MT) in the Owl Monkey (Aotus Trivirgatus). *Brain Res.* **81**, 199–213.

Allman, J. M. & Kaas, J. H. 1976 Representation of the visual field on the medial wall of occipital-parietal cortex in the Owl Monkey. *Science, N.Y.* **191**, 572–575.

Anderson, E. E. & Weymouth, F. W. 1923 Visual perception and the retinal mosaic. I. Retinal mean local sign – an explanation of the fineness of binocular perception of distance. *Am. J. Physiol.* **64**, 561–594.

Barlow, H. B. 1952 The size of ommatidia in apposition eyes. *J. exp. Biol.* **29**, 667–674.

Barlow, H. B. 1953 Summation and inhibition in the frog's retina. *J. Physiol., Lond.* **119**, 69–88.

Barlow, H. B. 1956 Retinal noise and absolute threshold. *J. opt. Soc. Am.* **46**, 634–639.

Barlow, H. B. 1957 Purkinje shift and retinal noise. *Nature, Lond.* **179**, 255–256.

Barlow, H. B. 1958 Intrinsic noise of cones. In *Visual problems of colour, National Physical Laboratory Symposium* no. 8, vol. 2, pp. 617–630. London: H.M. Stationery Office.

Barlow, H. B. 1964 The physical limits of visual discrimination. In *Photophysiology*, vol. 2 (ed. A. C. Giese), pp. 163–202. New York: Academic Press.

Barlow, H. B. 1978 The efficiency of detecting changes of density in random dot patterns. *Vision Res.* **18**, 637–650.

Barlow, H. B. 1979 Reconstructing the visual image in space and time. *Nature, Lond.* **279**, 189–190.

Barlow, H. B. 1980 The absolute efficiency of perceptual decisions. *Phil. Trans. R. Soc. Lond.* B **290**, 71–82.

Barlow, H. B., Blakemore, C. & Pettigrew, J. D. 1967 The neural mechanism of binocular depth discrimination. *J. Physiol., Lond.* **193**, 327–342.

Barlow, H. B. & Fatt, P. (eds) 1977 In *Vertebrate photoreception*, pp. 301–304 and 351–353. London: Academic Press.

Barlow, H. B. & Levick, W. R. 1965 The mechanism of directionally selective units in the rabbit's retina. *J. Physiol., Lond.* **178**, 477–504.

Barlow, H. B., Levick, W. R. & Yoon, M. 1971 Responses to single quanta of light in retinal ganglion cells of the cat. *Vision Res.* **11** (suppl. 3), 87–102.

Barlow, H. B. & Reeves, B. C. 1979 The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Res.* **19**, 783–793.

Baylor, D. A., Lamb, T. D. & Yau, K-W. 1979 Responses of retinal rods to single photons. *J. Physiol., Lond.* **288**, 613–634.

Baylor, D. A., Matthews, G. & Yau, K-W. 1980 Two components of electrical dark noise in retinal rod outer segments. *J. Physiol., Lond.* **309**, 591–621.

Blake, R. 1978 Spatial vision in the cat. In *Frontiers of visual science, Springer Ser. opt. Sci.*, vol. 8 (ed. S. J. Cool & E. L. Smith), pp. 209–219. New York: Springer-Verlag.

Blakemore, C. & Van Sluyters, R. C. 1975 Innate and environmental factors in the development of the kitten's visual cortex. *J. Physiol., Lond.* **248**, 663–716.

Bonds, A. B. 1979 Development of orientation tuning in the visual cortex of kittens. In *Developmental neurobiology of vision, NATO adv. Study Inst. Ser. A*, vol. 27 (ed. R. D. Freeman), pp. 31–41. New York and London: Plenum Press.

Braitenburg, V. 1973 *Gehirngespinste. Neuroanatomie für kybernetisch Interessierte.* Heidelberg: Springer-Verlag. (Trans. *On the texture of brains*, Heidelberg Science Library, 1977.)

Burr, D. C. 1975 A second binocular depth perception system. BSc. thesis, University of Western Australia.

Burr, D. C. 1979 Acuity for apparent vernier offset. *Vision Res.* **19**, 835–837.

Burr, D. C. 1981 Temporal summation of moving images by the human visual system. *Proc. R. Soc. Lond.* B **211**, 321–339.

Burr, D. C. & Ross, J. 1979 How does binocular delay give information about depth? *Vision Res.* **19**, 523–532.

Chow, K-L., Blum, J. S. & Blum, R. A. 1950 Cell ratios in the thalamo-cortical visual system of *Macaca mulatta. J. comp. Neurol.* **92**, 227–239.

Cowey, A. 1979 Cortical maps and visual perception: the Grindley Memorial Lecture. *Q. Jl exp. Psychol.* **31**, 1–17.

Craik, K. J. W. 1938 The effect of adaptation on differential brightness discrimination. *J. Physiol., Lond.* **92**, 406–421.

Crick, F. H. C., Marr, D. H. & Poggio, T. 1981 An information-processing approach to understanding the visual cortex. In *The cortex* (ed. F. O. Schmitt). Boston: M.I.T. Press. (In the press.)

Daniel, P. M. & Whitteridge, D. 1961 The representation of the visual field on the cerebral cortex in monkeys. *J. Physiol., Lond.* **159**, 203–221.

Derrington, A. M. & Fuchs, A. F. 1981 The development of spatial frequency selectivity in kitten striate cortex. *J. Physiol., Lond.* (In the press.)

Enroth-Cugell, C. & Robson, J. G. 1966 The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol., Lond.* **187**, 517–552.

Ferster, D. 1981 A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J. Physiol., Lond.* (In the press.)

Fisken, R. A., Garey, L. J. & Powell, T. P. S. 1973 Patterns of degeneration after intrinsic lesions of the visual cortex (area 17) of the monkey. *Brain Res.* **53**, 208–213.

Fox, R., Lehmkule, S. W. & Westendorf, D. H. 1976 Falcon visual acuity. *Science, N.Y.* **192**, 263–265.

Freeman, R. D. & Bradley, A. 1980 Monocularly deprived humans: nondeprived eye has supernormal vernier acuity. *J. Neurophysiol.* **43**, 1645–1653.

Gilbert, C. D. & Kelly, J. P. 1975 The projections of cells in different layers of the cat's visual cortex. *J. comp. Neurol.* **163**, 81–106.

Goodwin, A. W., Henry, G. H. & Bishop, P. O. 1975 Direction selectivity of simple striate cells: properties and mechanisms. *J. Neurophysiol.* **38**, 1500–1523.

Guzman, A. 1968 Decomposition of a visual scene into three-dimensional bodies. *Proc. Fall Joint Computer Conf. (San Francisco, 1968)*, vol. 33, pt I, pp. 291–304.

Hecht, S., Shlaer, S. & Pirenne, M. 1942 Energy, quanta, and vision. *J. gen. Physiol.* **25**, 819–840.

Horridge, G. A. 1975 Optical mechanisms of clear-zone eyes. In *The compound eye and vision in insects* (ed. G. A. Horridge), pp. 255–298. Oxford: Clarendon Press.

Horridge, G. A. 1980 Apposition eyes of large diurnal insects as organs adapted to seeing. *Proc. R. Soc. Lond.* B **207**, 287–309.

Hubel, D. H. & Wiesel, T. N. 1959 Receptive fields of single neurones in the cat's striate cortex. *J. Physiol., Lond.* **148**, 574–591.

Hubel, D. H. & Wiesel, T. N. 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol., Lond.* **160**, 106–154.

Hubel, D. H. & Wiesel, T. N. 1970a The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol., Lond.* **206**, 419–436.

Hubel, D. H. & Wiesel, T. N. 1970b Stereoscopic vision in the Macaque monkey: Cells sensitive to binocular depth in area 18 of the Macaque monkey cortex. *Nature, Lond.* **225**, 41–42.

Hubel, D. H. & Wiesel, T. N. 1972 Laminar and columnar distribution of geniculo-cortical fibers in the Macaque monkey. *J. comp. Neurol.* **146**, 421–450.

Hubel, D. H. & Wiesel, T. N. 1974 Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. comp. Neurol.* **158**, 307–318.

Hubel, D. H. & Wiesel, T. D. 1977 Ferrier Lecture: Functional architecture of macaque monkey cortex. *Proc. R. Soc. Lond.* B **198**, 1–59.

Imbert, M. 1979 Maturation of visual cortex with and without visual experience. In *Developmental neurobiology of vision, NATO adv. Study Inst. Ser. A*, vol. 29 (ed. R. D. Freeman), pp. 43–46. New York and London: Plenum Press.

Julesz, B. 1971 *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Kirschfeld, K. 1979 The visual system of the fly: physiological optics and functional anatomy as related to behaviour. In *The neurosciences: fourth study program* (ed. F. O. Schmidt & F. E. Worden), ch. 17, pp. 297–310. Cambridge, Massachusetts: M.I.T. Press.

Kuffler, S. W. 1953 Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.* **16**, 37–68.

Land, M. F. 1965 Image formation by a concave reflector in the eye of the scallop, Pecten Maximus. *J. Physiol., Lond.* **179**, 138–153.

Land, M. F. 1976 Superposition images are formed in the eyes of some oceanic decapods. *Nature, Lond.* **263**, 794–765.

Land, M. F. 1980 Compound eyes: old and new optical mechanisms. *Nature, Lond.* **287**, 681–686.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. 1959 What the frog's eye tells the frog's brain. *Proc. Inst. Radio Engrs* **47**, 1940–1945.

LeVay, S. & Gilbert, C. D. 1976 Laminar patterns of geniculocortical projection in the cat. *Brain Res.* **113**, 1–19.

Marshall, W. H. & Talbot, S. A. 1942 Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. *Biological symposia* (ed. J. Cattell), vol. 7, 115–164.

Michael, C. R. 1978a Color vision mechanisms in monkey striate cortex: simple cells with dual opponent-color receptive fields. *J. Neurophysiol.* **41**, 1233–1249.

Michael, C. R. 1978b Color sensitive complex cells in monkey striate cortex. *J. Neurophysiol.* **41**, 1250–1266.

Montero, V. M. 1980 Patterns of connections from the striate cortex to cortical visual areas in superior temporal sulcus of macaque and middle temporal gyrus of owl monkey. *J. comp. Neurol.* **189**, 49–59.

Morgan, M. 1975 Stereoillusion based on visual persistence. *Nature, Lond.* **256**, 639–640.

Morgan, M. J. 1976 Pulfrich effect and the filling in of apparent motion. *Perception* **5**, 187–195.

Morgan, M. J. & Thompson, P. 1975 Apparent motion and the Pulfrich effect. *Perception* **4**, 3–18.

Morgan, M. J. & Turnbull, D. F. 1978 Smooth eye tracking and the perception of motion in the absence of real movement. *Vision Res.* **18**, 1053–1059.

Nikara, T., Bishop, P. O. & Pettigrew, J. D. 1968 Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Expl Brain Res.* **6**, 353–372.

Peichl, L. & Wässle, H. 1979 Size, scatter and coverage of ganglion cell receptive field centres in the cat retina. *J. Physiol., Lond.* **291**, 117–141.

Platt, J. R. 1960 How we see straight lines. *Scient. Am.* **202**, 121–129.

Poggio, G. F. & Fischer, B. 1977 Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J. Neurophysiol.* **40**, 1392–1405.

Rauschecker, J. P. & Singer, W. 1981 The effects of early visual experience on the cat's visual cortex and their possible explanation by Hebb synapses. *J. Physiol., Lond.* **310**, 215–239.

Robson, J. G. 1975 Receptive fields: neural representation of the spatial and intensive attributes of the visual image. In *Seeing. Handbook of perception* (ed. E. C. Carterette & M. P. Friedman), vol. 5, pp. 81–116. New York: Academic Press.

Robson, J. G. 1980 Neural images: the physiological basis of spatial vision. In *Visual coding and adaptability* (ed. L. S. Harris, L. Erlbaum & associates), pp. 177–214. New Jersey: Hillsborough.

Robson, J. G. & Enroth-Cugell, C. 1978 Light distribution in the cat's retinal image. *Vision Res.* **18**, 159–173.

Ross, J. 1974 Stereopsis by binocular delay. *Nature, Lond.* **248**, 363–364.

Ross, J. & Hogben, J. H. 1975 The Pulfrich effect and short term memory in stereopsis. *Vision Res.* **15**, 1289–1290.

Sakitt, B. & Barlow, H. B. 1977 An economical encoding for size and position information. *Psychonomic Society Abstracts 1977*, Washington, D.C.

Sakitt, B. & Barlow, H. B. 1981 A model for the economical cortical encoding of the visual image. *Biol. Cybernet.* (In the press.)

Shlaer, R. 1972 An eagle's eye: quality of the retinal image. *Science, N.Y.* **176**, 920–922.

Sillito, A., Kemp, J. A., Milson, J. A. & Bernardi, N. 1980 A reevaluation of the mechanism underlying simple cell orientation selectivity. *Brain Res.* **194**, 517–520.

Snyder, A. W. 1972 Coupled-mode theory for optical fibers. *J. opt. Soc. Am.* **62**, 1267–1277.

Snyder, A. W. & Miller, W. H. 1977 Photoreceptor diameter and spacing for highest resolving power. *J. opt. Soc. Am.* **67**, 696–698.

Snyder, A. W. & Miller, W. H. 1978 Telephoto lens system in falconiform eyes. *Nature, Lond.* **275**, 127–129.

Tulunay-Keesey, U. 1960 Effects of involuntary eye movements on visual acuity. *J. opt. Soc. Am.* **50**, 769–774.

Tusa, R. J. & Palmer, L. A. 1980 Retinotopic organization of areas 20 and 21 in the cat. *J. comp. Neurol.* **193**, 147–164.

Van Essen, D. C. & Maunsell, J. H. R. 1980 Two-dimensional maps of the cerebral cortex. *J. comp. Neurol.* **191**, 255–281.

Van Meeteren, A. & Barlow, H. B. 1981 The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vision Res.* (In the press.)

Vogt, K. 1975 Zur Optik der Flusskrebsauges. *Z. Naturforsch.* **C30**, 691.

Walls, G. L. 1940 Postscript on image expansion by the foveal clivus. *Arch. Ophthal.* **23**, 831–832.

Walls, G. L. 1942 *The vertebrate eye and its adaptive radiation.* Bloomfield Hills, Michigan: Cranbrook Institute of Science. (Republished in 1963 by Hafner, New York.)

Wässle, H., Boycott, B. B. & Illing, R.-B. 1981 Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations. *Proc. R. Soc. Lond.* B **212.** (In the press.)

Westheimer, G. & McKee, S. P. 1975 Visual acuity in the presence of retinal image motion. *J. opt. Soc. Am.* **65**, 847–850.

Willmer, E. N. & Wright, W. D. 1945 Colour sensitivity of the fovea centralis. *Nature, Lond.* **165**, 119–121.

Wurtz, R. H. & Albano, J. E. 1980 Visual-motor function of the primate superior colliculus. *A. Rev. Neurosci.* **3**, 189–226.

Yau, K-W., Matthews, G. & Baylor, D. A. 1979 Thermal activation of the visual transduction mechanism in retinal rods. *Nature, Lond.* **279**, 806–807.

Zeeman, E. C. 1962 The topology of the brain and visual perception. In *Topology of 3-manifolds and related topics* (ed. M. K. Fort), pp. 240–256. Englewood Cliffs, New Jersey: Prentice Hall.

Zeki, S. 1974 Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol., Lond.* **236**, 549–573.

Zeki, S. 1978 Functional specialization in the visual cortex of the rhesus monkey. *Nature, Lond.* **274**, 423–428.

Zeki, S. 1980 The representation of colours in the cerebral cortex. *Nature, Lond.* **284**, 412–418.