Single units and sensation: A neuron doctrine for perceptual psychology?

H B Barlow
Department of Physiology-Anatomy, University of California, Berkeley, California 94720
Received 6 December 1972

Abstract. The problem discussed is the relationship between the firing of single neurons in sensory pathways and subjectively experienced sensations. The conclusions are formulated as the following five dogmas:
1. To understand nervous function one needs to look at interactions at a cellular level, rather than either a more macroscopic or microscopic level, because behaviour depends upon the organized pattern of these intercellular interactions.
2. The sensory system is organized to achieve as complete a representation of the sensory stimulus as possible with the minimum number of active neurons.
3. Trigger features of sensory neurons are matched to redundant patterns of stimulation by experience as well as by developmental processes.
4. Perception corresponds to the activity of a small selection from the very numerous high-level neurons, each of which corresponds to a pattern of external events of the order of complexity of the events symbolized by a word.
5. High impulse frequency in such neurons corresponds to high certainty that the trigger feature is present.
The development of the concepts leading up to these speculative dogmas, their experimental basis, and some of their limitations are discussed.

1 Introduction

In this article I shall discuss the difficult but challenging problem of the relation between our subjective perceptions and the activity of the nerve cells in our brains. Results obtained by recording from single neurons in sensory pathways have aroused a lot of interest and obviously tell us something important about how we sense the world around us; but what exactly have we been told? In order to probe this question, ideas that fit current knowledge as well as possible must be formulated, and they must be stated clearly enough to be tested to see if they are right or wrong; this is what I have tried to do. The central proposition is that our perceptions are caused by the activity of a rather small number of neurons selected from a very large population of predominantly silent cells. The activity of each single cell is thus an important perceptual event and it is thought to be related quite simply to our subjective experience. The subtlety and sensitivity of perception results from the mechanisms determining when a single cell becomes active, rather than from complex combinatorial rules of usage of nerve cells.

In order to avoid vagueness, I have formulated this notion in five definite propositions, or dogmas, and the reader who wishes to see the trend of this article can glance ahead (to page 380). Some of the dogmas will be readily accepted by most people who hope to find a scientific basis for human thought processes, but I felt they required statement and discussion in spite of their widespread tacit acceptance. Others are more original, will be challenged by many, and have the nature of extrapolations from the current trend of results rather than conclusions reasonably based upon them. Before these dogmas are stated the developments that have led to them will be briefly reviewed. The literature is extensive, and much of it will have been incorporated into the reader's common knowledge. My aim, therefore, is to pick out the conceptual turning points in order to show the direction we are headed.

After stating the dogmas, criticisms and alternatives will be discussed in an attempt both to justify them and to clarify them further.

2 Recording from single neurons

2.1 *Peripheral nerves*

In the twenties and thirties methods were developed for amplifying and recording the weak transient electrical potentials associated with the activity of nerve fibres, and Adrian and his colleagues used these methods to record the all-or-none impulses of single nerve fibres connecting the sense organs to the brain (Adrian, 1926a, 1926b; Adrian and Zotterman, 1926a, 1926b; Adrian, 1928). They showed, for example, that each fibre coming from the skin responded to a particular type of stimulus, such as pressure, temperature change, or damage, applied to a specific region or receptive field. The frequency of the impulses depended upon the intensity of the stimulus, but it was clear that the character of the sensation (touch, heat, or pain) depended upon the fibre carrying the message, not the nature of the message, since this consisted of trains of similar impulses in all fibres. Nerves had long been recognized as the link between physical stimulus and sensation, so these results provided physiological flesh and blood to the skeleton that anatomical studies had revealed a long time earlier.

Most of the results confirmed another ancient idea, namely Müller's doctrine of specific nerve energies: the specificity of different sensations stems from the responsiveness of different nerve fibres to different types of stimulus. The chemical senses proved to be a little different (Pfaffman, 1941, 1955; Ganchrow and Erickson, 1970), but in spite of the fact that they did not quite fall in line, the concept that resulted from two decades of recording from peripheral fibres and following their connections in the brain was of a simple mapping from sense organs to sensorium, so that a copy of physical events at the body surface was presented to the brain (Bard, 1938; Marshall *et al.*, 1941; Adrian, 1941, 1947). Some modification was recognized to occur, for sensory nerves usually adapt to a constant stimulus, and therefore signal sudden changes of stimulus energy better than sustained levels. Neighbouring receptive fields and modalities were also known to overlap, but when the activity of neurons at higher levels in sensory pathways was recorded it became obvious that something was happening more complex and significant than could be fitted into the concept of simple mapping with overlap and adaptation.

2.2 *Sensory neurons of the retina*

Starting with Granit (Granit and Svaetichin, 1939; Granit, 1947) and Hartline (1938; 1940a, 1940b) in the retina, and Galambos and Davis (Galambos and Davis, 1943; Galambos, 1944; Galambos and Davis, 1948) at the periphery of the pathway for hearing, a generation of physiologists has studied sensory neurons in the central nervous system; all this obviously cannot be reviewed here, but we shall concentrate on the results that expanded the conceptual frame built on the earlier work. Previously it was possible for physiologists to be satisfied with describing how the sense organs and their nerves present a picture of the external world to the brain, and they were happy to leave it to the psychologists to discuss what happened next; but these next things started to happen around the physiologist's micro-electrodes, and he has to join the discussion.

The realization that physiological experiments can answer questions of psychological interest first dawned on me personally when I was working on the frog's retina. A vigorous discharge can be evoked from retinal ganglion cells by stimulating the appropriate region of the retina—the ganglion cell's 'receptive field' (Hartline, 1940a); but if the surrounding region is simultaneously stimulated the response of the cell is diminished or completely abolished (Barlow, 1953). This phenomenon is

called lateral inhibition, or peripheral suppression, and such a physiological mechanism had already been postulated in order to account for simultaneous brightness and Mach bands (Mach, 1886; Fry, 1948). Thus the physiological experiment was really providing evidence in support of a psychological hypothesis.

The invasion of psychological territory did not stop at this point. If one explains the responsiveness of single ganglion cells in the frog's retina using hand-held targets, one finds that one particular type of ganglion cell is most effectively driven by something like a black disc subtending a degree or so moved rapidly to and fro within the unit's receptive field. This causes a vigorous discharge which can be maintained without much decrement as long as the movement is continued. Now, if the stimulus which is optimal for this class of cells is presented to intact frogs, the behavioural response is often dramatic: they turn towards the target and make repeated feeding responses consisting of a jump and snap. The selectivity of the retinal neurons, and the frog's reaction when they are selectively stimulated, suggest that they are 'bug detectors' (Barlow, 1953) performing a primitive but vitally important form of recognition.

This result makes one suddenly realize that a large part of the sensory machinery involved in a frog's feeding responses may actually reside in the retina rather than in mysterious 'centres' that would be too difficult to understand by physiological methods. The essential lock-like property resides in each member of a whole class of neurons, and allows the cell to discharge only to the appropriate key pattern of sensory stimulation. Lettvin *et al.* (1959) suggested that there were five different classes of cell in the frog, and Levick, Hill and I (Barlow *et al.*, 1964) found an even larger number of categories in the rabbit. We called these key patterns 'trigger features', and Maturana *et al.* (1960) emphasized another important aspect of the behaviour of these ganglion cells: a cell continues to respond to the same trigger feature in spite of changes in light intensity over many decades. The properties of the retina are such that a ganglion cell can, figuratively speaking, reach out and determine that something specific is happening in front of the eye. Light is the agent by which it does this, but it is the detailed pattern of the light that carries the information, and the overall level of illumination prevailing at the time is almost totally disregarded.

It is true that Ingle (1968, 1971), Grüsser and Grüsser–Cornehls (1968), and Ewert (1970) have shown that it is too simple to suppose that feeding automatically and inevitably follows the activation of a certain class of retinal ganglion cells by their trigger features; higher coordinating mechanisms are also involved. Just as light is only an intermediate agent allowing a retinal ganglion cell to detect its trigger feature, so these optic nerve impulses must doubtless be regarded as intermediate agents enabling the higher centres to perform their tasks. We shall proceed to discuss these problems, but we have gained two important concepts from the frog's retina: it transmits a map, not of the light intensities at each point of the image, but of the trigger features in the world before the eye, and its main function is not to transduce different luminance levels into different impulse frequencies, but to continue responding invariantly to the same external patterns despite changes of average luminance.

2.3 *Sensory neurons of the cerebral cortex*

The function of the visual area of the mammalian cerebral cortex is obviously more relevant to the problem of our own subjective perceptions than is the frog's retina, and Hubel and Wiesel (1959) early discovered examples of selectivity for pattern in the responsiveness of cells in the visual cortex of cats. They found that a light or dark line, or a dark–light border, was required to evoke a vigorous response even in the simplest first-order cells. Furthermore the stimulus had to be at a rather precise

orientation and position in the visual field and in addition it usually had to be moving, often in a specific direction. Hubel and Wiesel (1962) also made a distinction between these cells and other classes with more elaborate stimulus requirements, which they believed corresponded to cells at later stages of information processing. They called these 'complex' and 'hypercomplex' units, and showed that they had properties suggesting that the input to each was from the simpler category of cells. The fascination of this analysis depends to a large extent upon successfully following the way units become selective for more and more complex properties at each stage. Some doubts have been cast on their hierarchical scheme (Stone, 1972), but it certainly gave new insight into how higher levels of categorization are developed from lower levels.

As well as the hierarchical concept, this work provided evidence for a new type of invariance. In the cat, as in the frog, the retina is mainly responsible for ensuring that the message sent to the brain is not much perturbed by changes in ambient illumination. In the cortex Hubel and Wiesel (1962) found that some of the higher level neurons responded to the same trigger feature over a considerable range of positions. The modality specificity of peripheral neurons indicates how one can, for instance, detect warmth at any point on the body surface, and we now see that the organized pattern specificity of a set of cortical neurons can in the same way produce positional invariance for pattern perception. This was previously one of the great puzzles, and, although we certainly do not understand how recognition is invariant for position, size, and perspective transformations, at least a start has been made.

Later experiments have shown that the primary neurons of the visual cortex are more specific in one respect than Hubel and Wiesel originally thought. They showed that most neurons are fed by inputs from both eyes, and they emphasized that the dominance of ipsi- or contra-lateral eye varied from cell to cell. Now it can be shown that a binocular stimulus often has to be very precisely positioned in both eyes in order to evoke the most vigorous response (Barlow et al., 1967; Pettigrew et al., 1968), and a more important variable than dominance emerges from the exact relative positions in the two eyes. Consider what must happen when the eyes are converged on some point in front of the cat and appropriate visual stimuli are presented; it is easy to position this stimulus correctly for either eye by itself, but, if it is to be correctly positioned for both, it will have to be at some specific distance from the cat. When the precise positioning for different units is studied, it is found that this specific distance for optimal response varies in different units in the same cortex, and among units serving the same region of visual space. Conversely, the selection of units which are activated provides the cat with some information about the distances of the various stimulus objects. In uncovering this aspect of the pattern selectivity of sensory neurons we again get the sense that a central neuron is reaching out to discover something important about what is happening in the real objective world. One even wonders if the line and edge detectors of Hubel and Wiesel may not have, as their main function, the linking together of information about the same object in the two retinal images in order to determine the object's most important coordinate— its distance from the animal. At all events, as in the case of the frog's bug-detector, the importance of the information abstracted from the retinal images gives some insight into the purpose or direction of the physiological mechanisms.

Something is known about these first steps of information processing in the visual cortex; what about the later stages? Results suggesting greater and greater specificity of response requirements have been obtained, and a nice example is the unit described by Gross et al. (1972) in the infero-temporal cortex of macaques; this responded best to stimulation by a figure with many of the specific characteristics of a monkey's hand, and the requirements of one such unit are well documented.

Work in this area is not easy to repeat, for one can readily see that it is largely a matter of chance to find a trigger feature of this order of complexity. Also, the possibility that cells may retain to adulthood the modifiable properties of immature cells that will be described later makes the prospect of investigating the sensory association areas an intimidating one.

Cortical neurons receive selective excitatory and inhibitory inputs from other neurons and thereby possess selective responsiveness for some characteristics and invariances for changes in other characteristics. This seems to have the potentiality of being a powerful information processing system.

3 Single units and psychophysics
The neurophysiological discoveries outlined above of course made a deep impression on those investigating sensation psychophysically, but although there are many superficial points of contact it has not proved easy to link sensations securely to specific patterns of neurophysiological activity. The topics I have chosen are again ones which seem to have implications about how we conceptualize this neuropsychic relationship.

3.1 *Lateral inhibition and simultaneous contrast*
The relation between lateral inhibition in the retina and simultaneous contrast has already been mentioned, but there is a large gap between the physiological level and the subjective effects shown in textbook illustrations, and it is too big to be bridged by a single simple statement. It is quite easy to show that frog and cat retinal ganglion cells demonstrate relevant effects, since their antagonistic surrounds (Barlow, 1953; Kuffler, 1953) make their responses depend upon contrast rather than absolute luminance. Hence on-centre cells respond to spots we would call white, off-centre cells to spots we would call black, even when the so-called black spot has a higher luminance than the white spot (Barlow *et al.*, 1957). But subjective contrast effects also hold for conditions where one cannot make such easy comparisons, for instance at the centre of an area which is much too large to fill the centre of a retinal receptive field. Of course, one can postulate some 'filling in' process (Yarbus, 1965), but the necessity of introducing ad hoc assumptions makes many explanations of subjective effects in terms of single units unconvincing.

The concept that enables one to escape this difficulty is to concentrate on the informational flow rather than on the direct subjective-physiological comparison. Information discarded in a peripheral stage of processing cannot be accurately added back centrally, and in the present case it helps to talk about 'attenuating low spatial frequencies' instead of 'signalling spatial contrast'. To say that some of the low-frequency attenuation of the whole visual system is performed by the opposed centre-surround organisation of the retinal ganglion cell (Campbell and Green, 1965; Enroth-Cugell and Robson, 1966) is more accurate than to say that all simultaneous contrast effects originate there.

3.2 *Colour*
In the field of colour vision De Valois has looked for relationships between various psychophysically measurable aspects of colour and the properties of single unit responses recorded at the level of the lateral geniculate nucleus. The main results provided a startling confirmation of Hering's long-standing hypothesis about the reciprocal organisation of colour systems (Svaetichin and MacNicholl, 1958; De Valois, 1960; Hurvich and Jameson, 1960; Wagner *et al.*, 1960), but the details are important. He has been able to establish neuro-psychic parallels using what may be called the 'lower envelope' or 'most sensitive neuron' principle. A monkey's ability to discriminate hue and saturation (De Valois *et al.*, 1966, 1967) is very close

to what one would expect if the monkey only pays attention to the most sensitive
of the optic nerves conveying information about these qualities of the stimulus. Thus
the psychophysical performance follows the lower envelope of the performance of
individual fibres. It is particularly interesting to see that a continuous psychophysical
function, hue discrimination as a function of wavelength, is served by a different type
of neuron in different ranges; over the long wavelength range the red–green opponent
system was much more sensitive to wavelength shift, whereas the blue–yellow system
was more sensitive at short wavelengths.

This result again fits in with the concept that neurophysiology and sensation are
best linked by looking at the flow of information rather than simpler measures of
neuronal activity. For instance it might be suggested that sensation follows the
average neural activity, and it would be easy to justify this on the neurophysiological
grounds that post-synaptic potentials are usually additive. However, this oversimple
suggestion is proved false by the fact that psychophysical hue discrimination does not
follow the average response of the red–green and blue–yellow systems, but instead
follows the lower envelope. Now when two noisy channels are both conveying
information about a signal, the channel with the highest signal/noise ratio dominates
the situation; the low signal/noise ratio channel can be used to improve performance
slightly, but it is a very small contribution except where its signal/noise ratio is nearly
as high as that of the more sensitive channel. Thus the 'most-sensitive neuron'
principle again fits the concept that, to link neurophysiological activity and sensation,
one should look at the flow of information.

3.3 Touch
Another example is given by the work of Mountcastle and his colleagues (Talbot
et al., 1968), in which they studied the responses at a number of levels to vibratory
stimuli applied to the glabrous skin of the hand. First they recorded from cutaneous
afferents in the monkey, then the cortical responses in the same species, finally they
made psychophysical measures of sensory responses in humans to the same stimuli.
As with the work on colour, they established that the sensory response depends
simply upon the category of nerve fibre with the lowest threshold.

The fact that the subjective sensation in both the colour and touch systems seems
to follow the lower envelope of the responses of the various types of sensory neuron
may give an important clue to the way in which these neurons represent sensations.
It is as if the screen on which sensations appear is completely blank until a sensory
pathway is activated, but when this happens a point lights up and becomes instantly
visible. This is not what one would expect if there was a lot of ongoing activity in
all pathways, or if the magnitude of the signal was a linear function of intensity, nor
is it what one would expect if sensation depended in a complex combinatorial way
upon the activity of many units. Rather it suggest the concept that the magnitude of
the signal directly represents the signal/noise ratio, for then the insignificant signals
will automatically be small, and the neuron firing most will automatically be the
most sensitive. This concept receives some support in the next section and is taken
up in the fifth dogma and its discussion.

3.4 Adaptation after-effects
The fact that one is almost unaware of the constant pressure applied to the skin by
the chair one is sitting on presumably results, at least in part, from the rapid decline
in frequency of the volley of sensory impulses initiated by contact (Adrian, 1928).
Central neurons that respond to specific patterns of sensory input also give a
decreased response when the pattern is sustained or repeatedly presented, though
there have actually been surprisingly few investigations of this effect. These

adaptation, habituation, or fatigue effects lead to plausible explanations for many well-known sensory illusions.

For example the rate of discharge in the directionally selective neurons of the rabbit retina declines if a stimulus is continuously moved through the receptive field in the preferred direction, and following cessation of movement the maintained discharge is found to be suppressed (Barlow and Hill, 1963). The resulting imbalance between neurons signalling opposite directions seems to provide a ready explanation of the apparent reversed movement of stationary objects following prolonged inspection of moving objects (the so-called 'waterfall effect'), and provides another example of an ancient psychophysical hypothesis (Wohlgemuth, 1911) being confirmed neurophysiologically. One must bear in mind that these neural effects were described in the rabbit's retina, whereas in the human, as in the cat and monkey, neurons are probably not directionally selective until the level of the visual cortex (Barlow and Brindley, 1963), but the same type of explanation may well apply to neurons at this level.

It has been suggested that one can make inverse inferences from the existence of an after-effect to the presence of neurons with particular selective responses. This is no place to argue whether the after-effects of adaptation to gratings imply a Fourier-type analysis (Blakemore and Campbell, 1969), or whether they can be satisfactorily accounted for by families of different-sized neurons with conventional Hubel–Wiesel-type receptive fields, but there is certainly room for argument, and this makes selective adaptation a difficult tool to use to discover later stages of information processing.

Instead, I think the importance of sensory adaptational effects, and of the corresponding neurophysiological phenomena, lies in the support both these phenomena lend to the concept put forward at the end of the last section. If sensory messages are to be given a prominence proportional to their informational value, mechanisms must exist for reducing the magnitude of representation of patterns which are constantly present, and this is presumably the underlying rationale for adaptive effects.

3.5 Noisiness or reliability of single units
It used to be commonly held that nerve cells were unreliable elements, much perturbed by metabolic or other changes and perhaps also by random disturbances of more fundamental origin (McCulloch, 1959; Burns, 1968). The fairly high degree of reliability that the nervous system achieves as a whole was explained by the supposed redundancy of neural circuits and appropriate rules for averaging and combining them. Developments in the study of human vision at the absolute threshold and of the absolute sensitivity of retinal ganglion cells in the cat now indicate that nerve cells are not intrinsically unreliable and that noise often originates externally.

Signal detection theory has familiarized psychologists with the problem of detecting signals in the presence of noise (Tanner and Swets, 1954; Green and Swets, 1966), and I think the assumed prevalence of internally generated noise was a major reason why this was thought to be an important new approach. But psychophysical studies have actually shown that the senses and the brain can operate with astonishing intrinsic reliability. Noise may always be present, but to an amazing extent it originates outside the nervous system. This was originally implied by the results of Hecht et al. (1942) on the absolute threshold of vision; they showed that about 100 quanta at the cornea, leading to 10 or less absorptions in the retina, were sufficient to give a sensation of light. But their most revolutionary finding was that the frequency-of-seeing curve, describing the breadth of the threshold zone, is mainly accounted for by quantum fluctuations, not internal sloppiness or random variations of the threshold criterion as had previously been thought. That is not to say that

'intrinsic retinal noise' or 'dark light' is non-existent or unimportant, for it is probably the main factor determining how many quanta are required for reliable detection (Barlow, 1956). It now appears probable that this originates in the photoreceptors and, in some subjects at least, is low enough to allow the conscious detection of the sensation caused by absorption of a single quantum (Sakitt, 1972); similar sensations occur in the absence of light stimuli, but at a lower frequency. In addition, the subjects can apparently discriminate between the sensory messages resulting from 2, 3, 4, etc. quantal absorptions, each being detected progressively more clearly and reliably.

This psychophysical work shows that the human brain, acting as a whole, can distinguish between the disturbances caused by small numbers of quantal absorptions. These must of course originate from single molecular events in single cells, but possibly the disturbance is thereafter diffused through many cells and abstracted in some way from a redundant neural representation. It therefore becomes very interesting to go into the neurophysiology and find how the absorption of a few quanta is signalled.

A sensitive example of a retinal ganglion cell of the cat, with its associated bipolar cells, receptors, amacrine and horizontal cells, will give a readily detectable discharge of impulses to as few as 2 or 3 quanta of light absorbed in the retina (Barlow et al., 1971). Such a stimulus will give rise to an average of 5 to 10 extra impulses. Thus a single quantal absorption causes as many as 3 extra impulses, two quanta cause about 6 impulses, and so on. The addition of 3 impulses to the maintained discharge is detectable on average, though, like the absorption of a single quantum in the human, it cannot be reliably detected on a single trial. There is of course some intrinsic noise, as shown by the maintained discharge, but its level is extraordinarily low when one considers that a single ganglion cell is connected to more than 100 rods containing a total of some 10^{10} molecules of rhodopsin, each poised ready to signal the absorption of a quantum. The important point is that quantitative knowledge of the noise level and reliability of single retinal ganglion cells enables one to see that the performance of the whole visual system can be attributed to a single cell: averaging is not necessary.

Individual nerve cells were formerly thought to be unreliable, idiosyncratic, and incapable of performing complex tasks without acting in concert and thus overcoming their individual errors. This was quite wrong, and we now realise their apparently erratic behaviour was caused by our ignorance, not the neuron's incompetence. Thus we gain support from this neuropsychical comparison for the concept of a neuron as a reliable element capable of performing a responsible role in our mental life, though we need not of course go to the other extreme and assume that mental errors are never caused by malfunctioning, ill-educated, or noisy neurons.

4 Modifiability of cortical neurons
The most recent conceptual change about the neural basis of our sensations has arisen from a reinspection of the origin of the selective responsiveness of cortical neurons.

4.1 Evidence for modifiability
Hubel and Wiesel (1963) at first thought they had shown that the whole of the elaborate organization responsible for the selectivity of neurons in the primary visual cortex was developed solely under genetic control. They reported that they found cortical neurons with normal adult-type specificity of responsiveness in young kittens which had not opened their eyes, or which had been deprived of visual experience by suture of their eyelids. In later investigations (Wiesel and Hubel, 1963, 1965; Hubel and Wiesel, 1965) they found that abnormal visual experience, such as unilateral

eye-suture, or prevention of simultaneous usage of the eyes by alternating occlusion or surgically induced strabismus, caused the development of an abnormal population of cortical cells. In accordance with their earlier findings they attributed this to a disruption of the preformed organization, and they discovered the very important fact that abnormal experience only modifies the cortex if it occurs during a particular 'sensitive' period—about 3 to 12 weeks in cats (Hubel and Wiesel, 1970).

Recent developments have extended these seminal findings, but they lead to somewhat different conclusions about the relative importance of experience and genetic factors in determining the selectivity of cortical neurons. First it was shown that kittens brought up with the two eyes exposed to different stimuli, one to vertical stripes, the other to horizontal, had a corresponding orientation selectivity of the receptive fields connected to each eye (Hirsch and Spinelli, 1970, 1971). This was confirmed in kittens exposed only to vertically or horizontally striped environments; these had no neurons sensitive to horizontally or vertically oriented stimuli respectively (Blakemore and Cooper, 1970). Evidence has been obtained that cats raised with a vertical displacement of the images in one eye induced by prisms also have abnormal vertical disparities of the pairs of receptive fields of cortical neurons connected to both eyes (Shlaer, 1971). Again, the cortex of a kitten exposed only to bright dots, with no contours or edges, contained units of an abnormal type responding well to small spots of light and showing little of the customary preference for lines (Pettigrew and Freeman, forthcoming). Furthermore it appears that a very brief period of exposure, as little as an hour, can have very pronounced effects on the subsequent selectivity of neurons in the visual cortex (Blakemore and Mitchell, 1973).

Such results could still possibly have been explained by disruption of the innately-determined highly-specific connections that were originally thought to underlie response specificity, but a reexamination of the properties of cortical neurons of kittens with no visual experience shows that they do not actually have fully-developed adult-type specificity (Barlow and Pettigrew, 1971). This is certainly the case with regard to disparity selectivity and, although there is directional preference and may be some weak orientation selectivity, they are not as narrowly selective as adult cells (Pettigrew, forthcoming). The anatomy of the developing cortex shows that only a small fraction of the normal complement of synapses is present before the critical period, and it is hard to believe that the cells could have adult properties (Cragg, 1972). It will take more work to determine the limits within which the pattern selectivity of cortical neurons can be modified, but the results already make it impossible to believe Hubel and Wiesel's original claim that many cells of the visually inexperienced kitten have the full adult-type selectivity.

4.2 Type of modification caused

It is instructive to look at the way in which experience modifies selectivity. In all cases the cortex of animals whose visual experience has been modified lacks neurons selectively responsive to patterns of excitation which a normal animal receives, but which have been excluded by the experimental modification. Thus unilateral lid suture led to a cortex with very few neurons excitable from the lid-sutured eye; likewise, alternating occlusion or strabismus, which decreases the probability of simultaneous excitation of corresponding neurons in the two eyes, decreased the proportion of neurons responding to both eyes. The same is true of the kittens reared in striped environments, or with a vertically deviating prism over one eye, or in an environment with point sources but no lines; in all these cases the rule holds that neurons are found for patterns of excitation that occur in the modified environment, but normally occurring types of selectivity are rare or absent if the patterns they would respond to have not been experienced in the modified environment.

This rule seems to amount to a striking confirmation of the speculation (Barlow, 1960) that a prime function of sensory centres is to code efficiently the patterns of excitation that occur, thus developing a less redundant representation of the environment. Previous examples of redundancy-reducing codes could be explained as genetically determined features of neural connectivity, but the above discoveries are definite examples of a modified code developed in response to a modified environment.

If on this page we have begun the correct story for simple cells of area 17, one can see that a book has been opened with regard to the properties of cells higher in the hierarchy, which are presumably themselves experience dependent and are fed by information from these experience-dependent neurons at the lower cortical levels. Even a small degree of modifiability would be extraordinarily significant in a hierarchically organized system, just as, in evolution, weak selection pressure is effective over many generations.

5 Current concept of the single neuron

The cumulative effect of all the changes I have tried to outline above has been to make us realise that each single neuron can perform a much more complex and subtle task than had previously been thought. Neurons do not loosely and unreliably map the luminous intensities of the visual image onto our sensorium, but instead they detect pattern elements, discriminate the depth of objects, ignore irrelevant causes of variation, and are arranged in an intriguing hierarchy. Furthermore, there is evidence that they give prominence to what is informationally important, can respond with great reliability, and can have their pattern selectivity permanently modified by early visual experience. This amounts to a revolution in our outlook. It is now quite inappropriate to regard unit activity as a noisy indication of more basic and reliable processes involved in mental operations; instead, we must regard single neurons as the prime movers of these mechanisms. Thinking is brought about by neurons, and we should not use phrases like 'unit activity reflects, reveals, or monitors thought processes', because the activities of neurons, quite simply, *are* thought processes.

This revolution stemmed from physiological work and makes us realize that the activity of each single neuron may play a significant role in perception. I think that more clearly stated hypotheses are now needed about these roles in order to allow our psychological knowledge and intuitions about our perceptions to help us plan future experiments.

6 Five propositions

The following five brief statements are intended to define which aspect of the brain's activity is important for understanding its main function, to suggest the way that single neurons represent what is going on around us, and to say how this is related to our subjective experience. The statements are dogmatic and incautious because it is important that they should be clear and testable.

6.1 *First dogma*

A description of that activity of a single nerve cell which is transmitted to and influences other nerve cells, and of a nerve cell's response to such influences from other cells, is a complete enough description for functional understanding of the nervous system. There is nothing else 'looking at' or controlling this activity, which must therefore provide a basis for understanding how the brain controls behaviour.

6.2 *Second dogma*

At progressively higher levels in sensory pathways information about the physical stimulus is carried by progressively fewer active neurons. The sensory system is

organized to achieve as complete a representation as possible with the minimum number of active neurons.

6.3 *Third dogma*
Trigger features of neurons are matched to the redundant features of sensory stimulation in order to achieve greater completeness and economy of representation. This selective responsiveness is determined by the sensory stimulation to which neurons have been exposed, as well as by genetic factors operating during development.

6.4 *Fourth dogma*
Just as physical stimuli directly cause receptors to initiate neural activity, so the active high-level neurons directly and simply cause the elements of our perception.

6.5 *Fifth dogma*
The frequency of neural impulses codes subjective certainty: a high impulse frequency in a given neuron corresponds to a high degree of confidence that the cause of the percept is present in the external world.

7 First dogma: Significant level of description
This dogma asserts that a picture of how the brain works, and in particular how it processes and represents sensory information, can be built up from knowledge of the interactions of individual cells. At the moment single-unit electrical recording is the only tool with temporal and spatial resolution adequate to locate the effect of a particular sensory stimulus in a particular cell. Other tools (biochemical, electron microscopy, etc.) can obviously provide essential information about these interactions, but the dogma may be criticized more fundamentally; it may be suggested that the whole problem should be approached at a different level. One could attack from either side, suggesting either that one should look at grosser signs of nervous activity, such as the weak extracellular potentials that result from the activity of many neurons, or that one should approach the problem at a more microscopic level, studying synaptic and molecular changes.

Interest in evoked potentials and electroencephalography has waned partly because their study led to slow progress compared with single-unit recording, but also because the rationale for their use was undermined. A prime reason for attending to these macroscopic manifestations of nervous activity was the belief that individual cells were too unreliable to be worthy of attention singly, and hence it was better to look at a sign of activity that resulted from many of them. Here, it was thought, may be a property of a group of cells analogous to temperature or pressure as a property of a collection of molecules that individually behave randomly. The demonstration that single nerve cells have diverse and highly specific responsiveness to sensory stimuli, and are astonishingly reliable, showed the fallacy of this analogy.

The search for a molar property of a mass of working nerve cells is certainly not worthless. Physiologists, and all biologists for that matter, tend to be emotionally divided into globalists and atomists. The globalists are amazed at the perfection of functioning of the whole animal, and they observe that the atomists' analytical investigations of living matter always leave unexplained many of the most remarkable attributes of the intact animal. As a result the globalist can play a crucially important role in pointing out where the atomists' explanations are incomplete. Now the brain does much more interesting things than produce weak extracellular potentials: it controls behaviour, and this is surely the global product that, at our present state of understanding, really does appear greater than the sum of its parts. It would be no use looking at single neurons if it will be forever impossible to explain overall behaviour in terms of the actions and interactions of these subunits;

if that were so, the globalists' despair would be justified. On the other hand it is precisely because rapid progress has been made that this article is being written; it no longer seems completely unrealistic to attempt to understand perception at the atomic single-unit level.

The second criticism, that one should approach the problem at a more microscopic level, is really only answerable by saying, "*Go ahead and do it*", for undoubtedly there is much to be learned at a synaptic and molecular level. But the important question here is whether lack of this knowledge will impede a major advance in our conception of how the brain works. The dogma asserts that it is the intercellular actions and interactions that possess the elaborate organization responsible for behaviour; hence it asserts that knowledge at a more-microscopic intracellular level is not a prerequisite for understanding such organization.

8 Second dogma: The economical representation of sensory messages
The main task in this section is to discern the principles that underlie the changes in characteristic responsiveness of single units at successive levels in sensory pathways. The aim is to understand how sensory information is represented or 'displayed'. The successive levels to be considered will be peripheral photoreceptors and cutaneous afferents; retinal ganglion cells of the cat, frog, or rabbit, the latter of which seem to exemplify a more complex type of processing; and the visual cortex of cats. Obviously these are not an ideal series for comparisons and extrapolations, but they are the best we can do.

The discussion initially revolves around three issues: changes in the degree of specificity and generality of the stimuli to which the cells respond; changes in the number of parallel categories of selectively sensitive cells that carry the information; and changes in the number of the cells that one may expect to be activated by normal visual scenes. What emerges is that, at the higher levels, fewer and fewer cells are active, but each represents a more and more specific happening in the sensory environment.

8.1 *Specificity and generality of responsiveness*
The pattern specificity of sensory neurons is the aspect that is most widely emphasized: it was spectacular to discover single neurons in the retina responding to movement of the image in a specific direction, cortical neurons responding only to slits of light at a particular orientation, and a unit in the infero-temporal cortex that responds best to a monkey's hand. But the invariance of the response to changes in the stimulus is equally remarkable. A retinal unit continues to respond to direction of motion in spite of many decades of change in input luminance or contrast, in fact in spite of reversal of contrast (Barlow, 1969a). At the cortical level a complex cell insists that a stimulus is appropriately oriented, but will respond in spite of wide variations of position (Hubel and Wiesel, 1962). And the monkey-paw unit similarly retains its pattern specificity over a large part of the visual field (Gross *et al.*, 1972).

In talking about these properties of sensory neurons actual examples are perhaps more informative than the words specific and general. A single receptor containing a red-sensitive pigment is specific in the sense that long-wavelength light must be present at a particular part of the image in order to excite it, and it is general in the sense that all images with this property will excite it. In contrast to this type of specificity and generality, the high-level neurons are no longer limited to purely local attributes of the image. They are selective for pattern, which requires that a considerable region of the image is taken into account. But there are other aspects of their specific selectivity that also need to be considered.

8.2 *Number of selective categories*

At the level of receptors there are a small number of different sensory modalities picking up, in parallel, information from different positions. This is the case both for the half dozen types of cutaneous sensation, and for the smaller number of retinal receptor types responding to the visual image. At the level of ganglion cells the number of sub-modalities or selective categories has greatly increased. Consider the rabbit, where there are the following (Barlow *et al.*, 1964; Levick, 1967): two concentric types (on- and off-centre); four on–off type directionally selective (for movements up, down, antero–posterior, and postero–anterior); three directions for slow, on-type, directionally selective; one type sensitive to fast movement; one type sensitive to 'uniformity'; and, confined to the visual streak, two types of orientation-selective neurons, neurons selective for slow-moving small objects, and neurons selective for edges. This makes a total of 15 different selective categories. In addition there must be units signalling colour, since the rabbit shows behavioural evidence for it, but these have not yet been found in the retina.

Now move up to the simple cells in area 17 of cat cortex. These vary in position, orientation, disparity, and size, as well as being selective for light bars, dark bars, or edges. The evidence is not sufficient to say how many distinct selective categories these form, but for each of the first three variables the resolution of a single neuron is good, in the sense that small departures from the preferred position, orientation, or disparity cause large decreases of response amplitude (Bishop, 1970). These variables already define four dimensions, and we have not yet considered size specificity, velocity specificity, nor the additional complexities of light, dark, or edge detectors, and of course colour. There are certainly several orders of magnitude more neurons in the primary projection area than there are input fibres, or resolvable points in the visual field, and it is abundantly clear that the number of selective categories has increased enormously. Activity of a particular neuron signifies much more than the presence of light at a particular locus in the visual field; its activity signifies a great deal about the nature of the pattern of light at that locus.

The fact that many parallel communication channels are used in the nervous system has been widely recognised, but here we see an enormous expansion of the number of parallel paths, and this occurs without much redundant reduplication of channels, for each neuron seems to have a different specific responsiveness. It is as if, at high levels, the size of the alphabet available for representing a sensory message was enormously increased. Perhaps it would be better to say that, if the activity of a low-level neuron is like the occurrence of a letter, that of a high-level neuron is like the occurrence of a word—a meaningful combination of letters. But to understand this better we must look at the third aspect of the way sensory messages are represented at different levels, namely the proportion of neurons that are usually active. If the pattern of activity caused by a visual scene has, on average, K neurons active out of the total of N neurons, then we have seen above that N increases at high levels; can one say anything about how K changes?

8.3 *Number of active cells*

If one considers the retinal cones under typical photopic conditions, the vast majority must be partially active. They may be nearer the depolarized than the hyperpolarized limit of their dynamic range, but the majority will be somewhere well within it. For the retinal ganglion cells of a cat the situation is a little different; while a few units, those corresponding to the brightest and dimmest parts of the scene, will be vigorously active, the majority, corresponding to the parts of the scene near the mean luminance, will be discharging at rates close to their maintained discharge level, which in its turn is near the low-frequency end of their dynamic range. Thus there will be a lot of units with low degrees of activity and a few which

are vigorously active. Recoding in the retina changes the distribution of activity so that low impulse frequencies are common, high impulse frequencies rare.

Now consider the rabbit, with its more elaborate retinal processing, and greater richness of pattern–feature signalling neurons. It is characteristic of the more specific of these neurons that they have a very low maintained discharge, and are extremely hard to excite until their exact trigger feature has been found. One flashes lights, waves wands, and jiggles 'noise figures' in the appropriate part of the visual field for many minutes, maybe hours, before finding the right combination for excitation. It is reasonably certain that the right combination does not occur often in the natural environment either, and therefore these units must spend only a small fraction of the time in an active state. Low impulse frequencies are even commoner, high impulse frequencies even rarer, than in cat retina.

For the cat cortex this trend is carried further, and one can see another aspect emerging. If one takes a small region of the visual field, it either does contain a bright bar, dark bar, or edge, or, much more likely, it does not. Thus, like the rabbit units, the cells with these specific responsivities must be only infrequently active. But in addition, on the rare occasions when one of the appropriate trigger features is present, it is one of a set which tend to be mutually exclusive: a bright bar cannot be a dark bar, and it can have only one orientation and disparity. The stimulus selects which cell to activate from a range of many possible cells, and it is pretty well impossible to activate simultaneously more than a small fraction of this number.

The picture developing is that at low levels visual information is carried by the pattern of joint activity of many elements, whereas at the upper levels of the hierarchy a relatively small proportion are active, and each of these says a lot when it is active. But, although we clearly see that the proportion active, K/N, decreases, we cannot tell whether it decreases as rapidly as N increases, and thus we still do not know how K itself changes. The second dogma goes beyond the evidence, but it attempts to make sense out of it. It asserts that the overall direction or aim of information processing in higher sensory centres is to represent the input as completely as possible by activity in as few neurons as possible (Barlow, 1961, 1969b). In other words, not only the proportion but also the actual number of active neurons, K, is reduced, while as much information as possible about the input is preserved.

By now much can one reasonably expect K to be reduced? One requires the concepts of channel capacity and redundancy from information theory (Shannon and Weaver, 1949; Woodward, 1953) to make a rough estimate. Some reduction can be accomplished without any loss of information simply by the increase of N. K/N is the probability of a fibre being active, and, if it is the same for all neurons, the information capacity of a set of N neurons, either active or not active, is $-K\log_2(K/N) - (N-K)\log_2[(N-K)/N]$. If K/N is small, the second term contributes little; the capacity then is, approximately, the number of active neurons times the information provided by each active neuron, and this increases directly as the negative logarithm of the probability of it being active, $-\log(K/N)$. Hence the number active can be reduced as N increases without loss of information capacity, but by itself this does not allow K to be reduced very much: for instance, if we suppose that $\frac{1}{4}$ of the 2×10^6 optic nerve fibres are active and that there are 10^8 cortical neurons receiving this information, then one finds that $1 \cdot 5 \times 10^5$ cortical neurons must, on average, be active in order to have the same information capacity as the 5×10^5 active optic nerve fibres. But this applies only to capacity, and a substantial reduction in K is possible on the basis of another principle.

Visual information is enormously redundant, and it has been suggested previously that sensory coding is largely concerned with exploiting this redundancy to obtain

more economical representation of the information. If the argument is correct, the number of active neurons can be reduced, but it is very difficult even to guess how big a reduction in K such recoding can achieve; if it is $\frac{1}{10}$ up to the cortex, and another factor of $\frac{1}{10}$ achieved in visual I, II and III, one would end up with about 1000 active fibres carrying the information provided by 5×10^5 active optic nerve fibres; though the reductions might be substantially greater or less, this is the order of magnitude of the reduction contemplated.

According to dogma, these 1000 active neurons represent the visual scene, but it is obvious that each neuron must convey an enormously larger share of the picture than, say, one point out of the quarter million points of a television picture. Perhaps a better analogy is to recall the 1000 words that a picture is proverbially worth; apparently an active neuron says something of the order of complexity of a word. It seems to me not unreasonable to suppose that a single visual scene can be represented quite completely by about 1000 of such entities, bearing in mind that each one is selected from a vast vocabulary and will in addition carry some positional information.

9 Third dogma: Selectivity adapted to environment
9.1 *Evolutionary adaptation*
Some economies of the type indicated above can be achieved by exploiting forms of redundancy which are present in all normal environments. Levels of sensory stimulation do not range at random over the whole scale of possible values, and it makes sense to regard adaptation of peripheral receptors as a measure to achieve economy by signalling changes from the mean instead of absolute values. Similarly in most situations neighbouring points on a sensory surface are more likely to be similar than distant points, and it thus makes sense to regard contrast enhancement by lateral inhibition as another economy measure. The argument can be carried on to cover the redundancy-reducing value of movement, edge, or disparity detectors (Barlow, 1969b), but, if these are genetically-determined redundancy-reducing codes, they must be fixed once and for all during development, and they could only work for redundant properties of all sensory environments. The hypothesis becomes more interesting when one considers the possible mechanisms for achieving economy by exploiting the redundancy of particular sensory scenes, for this requires storage of information and plasticity of the neural structures involved.

9.2 *Reversible adaptation*
The neural changes of dark and light adaptation may be regarded as a simple example of reversible plasticity achieving this end. The luminance corresponding to zero impulses is affected by the past history of illumination and by the surrounding luminances in such a way that the majority of fibres are responding at low frequencies. But, even though this involves definite changes in the synaptic transfer properties of retinal neurons, the statistical characteristic of visual images that enables this to achieve economy is always the same, namely the fact that the distribution of luminances is grouped around local and temporal mean values, so that small deviations from the mean are commoner than large deviations (Barlow, 1969a). Hence the most commonly occurring luminances require fewest impulses.

9.3 *Permanent adaptation*
The effects permanently impressed on the visual system during the sensitive period are the first example of plasticity for a particular type of redundancy. The distribution of orientational selectivity of primary neurons is biased in favour of the orientations the individual experienced during this critical time. If the analogy of a neuron's signal resembling the utterance of a word is recalled, this result suggests that the kitten's cortex only develops words for what it has seen. This could be brought

about by either selection or modification: are the dictionary words there, only the ones experienced becoming permanently connected; or do the cells themselves determine that a frequently experienced pattern, such as lines of a particular range of orientations, are events for which words are desirable? The evidence favours modification, and the idea to which it leads of the successive hierarchical construction of a dictionary of meaningful neurons has enormous appeal. For the present we can only justify the third dogma by saying the evidence suggests such a dictionary may be built up, though we are far from being able to look into its pages by physiological methods. In the next section we turn to the subjective view of this dictionary.

10 Fourth dogma: Origin of perceptions
10.1 *Personal perception*
In order to delimit more accurately what this dogma does and does not say it may be useful to define and separate three mysteries of perception. The first is the personal, subjective, aspect of my experience of, say, the red pencil with a blue eraser in my hand. There does not seem to be anything that could be said about the activity of nerve cells accompanying this experience that would in any way 'explain' the aspect of it that is mysterious, personal, and subjective. I think this part of the experience is something that one must be content to leave on one side for the moment, but it is important that this part of subjective experience almost always accompanies electrical stimulation of a peripheral sensory nerve, and usually accompanies electrical stimulation of the sensory areas of the brain, for this implies that the full subjective experience, including this mysterious personal element, accompanies the neural events of sensation, however these are caused. This fact strongly suggests that it is no waste of time to look into these neural events: beauty is a mysterious attribute of a work of art, but that does not imply that you cannot create a beautiful painting by non-mysterious material means.

10.2 *Conscious perception*
The second mystery is that we are not consciously aware of much that goes on in our brains, so the inverse of the fourth dogma is certainly not true: not every cortical neuron's activity has a simple perceptual correlate. Even at high mental levels much neural business is conducted without conscious awareness, and my own belief is that the conscious part is confined to experiences one communicates to other people, or experiences one is contemplating communicating to other people. This immediately introduces a social element into individual consciousness, for communication is impossible without a channel being open to a recipient. However, for present purposes we need only point out that interesting aspects of consciousness of this sort are by no means incompatible with the fourth dogma. An element of perception can possess a simple neural cause without it necessarily being the case that all simple neural events cause perception. There is therefore plenty of room for social, historical, or moral influences on perception, because these can influence the selection of the neural events that enter conscious perception.

10.3 *Validity of perceptions*
The third mystery about perceptions is why they are generally 'true': why are they so extraordinarily useful in guiding our actions and helping us to make decisions? This is the aspect that the second and third dogmas help one to understand. The economical and fairly complete representation of visual scenes by a reasonably small number of active neurons makes it much easier to visualize how they can be used for these purposes. The key point is that the active neurons carry the bulk of the information, and the vast number of inactive ones need not be taken into consideration. The difficulty of detecting among our sense impressions the entities

we use for rational thought has always been baffling: 'water', 'men', 'sheep', and even the simple letter 'A' represent particular logical functions of activity among the sensory neurons, but the number of possible logical functions is so vast that we are mystified how particular ones are realized, or why particular ones are selected for realization. The representation suggested by the second and third dogmas would allow relatively simple logical combinations to have properties approaching those required for the literal symbols of Boole (1854), the subjects of our conceptions. By using such symbols together with operational signs he founded mathematical logic, but the title of his major work, *"The Laws of Thought"*, clearly states his claim that his inquiries had "probable intimations concerning the nature and constitution of the human mind". It is gratifying to approach closer to an intuitively plausible neural realization of what he symbolized.

The notion that what we sense is a point by point representation of the physical signals impinging on our body has been rejected for psychological and philosophical reasons (see Boring, 1942), and more recent physiological evidence clearly supports this rejection. But this same evidence suggests that it should be replaced, not by return to a subjectively constructed phenomenology (Dreyfus, 1972), nor by the notion that we sense the world in terms of rigidly preordained 'structures', but by the deeper and more adaptive ideas of dogmas two and three; our sensorium is presented with a fairly small number of communications, each representing the occurrence of a group of external events having a word-like order of complexity, and, like words, having the special property that they lead to an economical representation of these physical events.

11 Fifth dogma: Signalling subjective certainty

There is one way in which the properties of neurons do not match up to the way Boole used symbols, for he insisted on their binary nature, or 'duality'. This is the property Aristotle called the principle of contradiction, and without it Boole's symbolic representation of logic would have been impossible. In contrast to this duality the response of a sensory nerve cell to its trigger feature consists of a volley of impulses lasting $\frac{1}{10}$ to 1 second, during which time the neuron can discharge any number of impulses between zero and nearly 1000. Therefore the response is graded, and it is not legitimate to consider it as a Boolean binary variable. The essential notion expressed in the fifth dogma is that a neuron stands for an idealization of reality whose complement can be formulated as a null hypothesis, and it is this that has the required Boolean logical property of duality. The idealizations should not be thought of as Kantian or Platonic, but rather as abstractions that model reality in the manner suggested by Craik (1943). The ideal populations of a statistician come even closer, for the parameters of such distributions model reality, and they are used to calculate whether or not a particular sample belongs within it. The process of idealizing the complement of the trigger feature will be clarified by a simple example.

Suppose we have a sensory neuron whose trigger feature is a simple physical event, such as the increase of light intensity at a specific position in the visual field. We are examining the suggestion that the graded responses this neuron gives to visual stimuli represent some function of the degree of certainty that the light did in fact increase, estimated from the physical events accessible to the neuron. For simplicity assume that a record is available of the total numbers of quanta absorbed in the receptive field of the sensory neuron during successive periods of about $\frac{1}{10}$ second duration up to and including the period about which it is to signal centrally. It is in principle possible to calculate the probability of occurrence of the observed number of quantal absorptions on the hypothesis that there was no change in the light intensity, and this is the test a statistician would apply to determine whether or not

the trigger feature was present. On this view impulse frequency signals some function of the significance level of a test of this sort, low probabilities corresponding to high impulse frequencies. Notice that the trigger feature is 'an increase of light', the idealization is 'there was no increase', and this idealization is based on observation of what has recently happened and therefore incorporates a model of the recent past.

The responses actually obtained to varying intensities of incremental stimulus fit quite well into this scheme (Barlow, 1969a), as do some less obvious features. The low-frequency maintained discharge could well represent the results, of low significance, obtained by testing the null hypothesis when no stimulus has been applied. Although individual fibres would not reach significance, information about quantal absorptions would be retained, and changes insignificant singly could be combined centrally to reach significance.

In the above example a high value of P and low impulse frequency would result from a shadow falling on the receptive field. The detection of this shadow might be of great survival significance for the animal, but the on-centre unit's trigger feature, null hypothesis, and statistical tests would be a poor way of detecting and signalling this important event. A different type of unit is required whose null hypothesis should be "There has been no *decrease* in the quantal absorption rate"; these would fire when the hypothesis is disproved by the quantal absorption rate dropping below the normal range of variation. Obviously these are the off-centre units, and it seems that the existence of complementary 'on' and 'off' systems fits the notion quite well. When there are a large number of neurons with trigger features that cannot coexist, as in area 17, these correspond to a large number of mutually exclusive hypotheses to be tested.

The fifth dogma clearly requires more development and testing, but it provides a possible answer to the question "What variable corresponds to impulse frequency in a high level sensory neuron?" Furthermore the answer ties it to a rather definitely felt subjective quantity—the sense of certainty.

12 Criticisms and alternatives
Single-unit recording hints at this probabilistic, adaptive, many-levelled, system for processing and displaying sensory information, but can we believe that what we perceive is the activity of a relatively small selection of upper-level units of this hierarchy? This is certainly a big jump beyond the present physiological evidence. We do not know how perspective transformations are disregarded, enabling us to perceive the same object irrespective of our angle of view, nor do we understand the mechanisms underlying size constancy, yet these mechanisms must intervene between the highest neurons we know about and quite simple perceptions. I think we have seen enough of what can be achieved in a few stages of neural image processing to believe that a few more stages could reach the point where a single neuron embodies, by virtue of its peripheral connections and their properties, an elementary percept, but let us examine an alternative and the evidence adduced in its support.

12.1 Combinatorial or holographic representations
The key suggestion about the organization of sensory processing that the second dogma asserts is that the information is carried by progressively fewer active neurons at progressively higher levels in the hierarchy. The brain receives complex patterns of activity in nerve fibres from the sensory receptors, and it generates complex patterns of outgoing commands to the muscles. It could be held that the patterns are equally complex at all the intervening stages as well, and this would mean that the significance of a single unit's activity would be virtually undecipherable without knowing what was going on in a host of other units. Certainly one would make

little progress in understanding a computer's operation by following the status of a single bit in its central processor, so this criticism is partly met by pointing to the success that has been achieved in the visual system by looking at the activity of units singly, one at a time. But we should also look critically at the main evidence advanced in favour of the combinatorial or holographic scheme.

12.2 *Mass action and resistance to damage*
The main argument that has been levelled against the view that individual cells play an important role in perception and in favour of a holographic representation is the reported fact that large parts of the cortex can be damaged with only minor resultant changes in behaviour or learning (Lashley, 1929, 1950). This led to Lashley's doctrine of 'cerebral mass action', but repetition of the original experiments and refinements in methods of testing, some of it by Lashley himself, have considerably weakened the original evidence in its favour (Zangwill, 1961). However, it certainly is remarkable that a mechanism with as much interdependence between its parts as the brain can function at all after it has been extensively damaged. A computer would not usually survive brain surgery or gunshot wounds, and it is therefore worthwhile considering the implications of the fact that the cortex is relatively immune to quite extensive injury.

The whole of a visual scene can be reconstructed from a small part of a hologram, with only slight loss of resolution and degradation of signal/noise ratio (Gabor *et al.*, 1971), so it has been claimed that the cortex must operate by some analogous principle in order to account for its resistance to damage. What is not widely appreciated is the fact that holography differs from ordinary image-recording photography not only in principle, but also in the materials used, for it requires photographic emulsions with resolutions of the order of the wavelength of light (Gabor *et al.*, 1971). With such materials a good quality 35 mm picture could easily be reproduced and repeated in every 1 mm^2 of the plate, and a plate containing such a reduplicated image would have to be pulverized into tiny pieces to prevent reconstructibility of the original from every fragment. The mass-action-like resistance to damage of the hologram is partly due to the enormous informational capacity of the materials that are required; immunity to damage is easy to achieve when such high redundancy is permissible, and this argument carries little weight in favour of holographic views of nervous operation.

Codes can be given error-correcting properties much less wastefully, and the argument can be turned around to favour the representation hypothesized in dogmas two and three. Because the few active cells have a fixed significance, and because the inactive ones are thought to carry so little information that they can be neglected, the only result of removing part of the cortex would be to eliminate some of the active neurons, and hence some of the perceptual entities, when a given scene is viewed. The 'meaning' of the active units in the undamaged cortex would remain the same, and might provide a sufficient basis for decision and action. This is very different from the situation where a neuron's activity has totally different significance depending upon the pattern of activity of which it forms a part, for if any of this pattern was in a damaged region the significance of activity in the undamaged part would be altered. Hence damage immunity is really an argument for neurons having an invariant meaning not dependent upon the activity of other neurons.

It should also be pointed out that very limited replication of 'percept neurons' would give considerable damage immunity: if a given neuron is replicated half a dozen times in different cortical regions there is a good chance of at least one of them surviving an extensive cortical ablation. This sixfold redundancy is enormously less than the holographic scheme possesses, and it can be concluded that the mass

action argument rebounds against the extensive combinatorial usage of neurons and actually favours the hypothesis of this article.

12.3 *Pontifical cells*

Sherrington (1941) introduced the notion of "one ultimate pontifical nerve-cell, ... the climax of the whole system of integration" and immediately rejected the idea in favour of the concept of mind as a "million-fold democracy whose each unit is a cell". Those who like the notion of perception as a cooperative or emergent property of many cells dismiss the suggestion that the activity of a single neuron can be an important element of perception by saying that, carried to its logical conclusion, it implies there must be a single 'pontifical cell' corresponding to each and every recognizable object or scene. First, notice that the current proposal does not say that each distinct perception corresponds to a different neuron being active, if perception is taken to mean the whole of what is perceived at any one moment; it says there is a simple correspondence between the elements of perception and unit activity. Thus the whole of subjective experience at any one time must correspond to a specific combination of active cells, and the 'pontifical cell' should be replaced by a number of 'cardinal cells'. Among the many cardinals only a few speak at once; each makes a complicated statement, but not, of course, as complicated as that of the pontif if he were to express the whole of perception in one utterance.

Two important difficulties arise from the notion of pontifical cells; first, if a separate neuron is needed for each of our perceptions, there are not enough to account for their almost incredible variety; second, the activity of a single isolated element would not convey anything of a perception's great richness, the connection between one perception and others. The 'grandmother cell' might respond to all views of grandmother's face, but how would that indicate that it shares features in common with other human faces, and that, on a particular occasion, it occurs in a specific position surrounded by other recognizable objects? Our perceptions simply do not have the property of being isolated unique events as one would expect if each corresponded to the firing of a unique neuron. Instead, they overlap with each other, sharing parts which continue unchanged from one moment to another, or recur at later moments in different contexts. I think the 'cardinal cell' representation surmounts these problems without any difficulty; if a critic can say how many different perceptions we are capable of, and how rich a network of relatedness exists between these perceptions, then one might be able to estimate how many cardinals' voices were required to represent these perceptions. But there is a misleading feature of the ecclesiastical analogy.

Most organizational hierarchies are pyramids: there are many members of the church, fewer priests, only a select number of cardinals, and a single pope. The hierarchy of sensory neurons is very different. It is true that there are more retinal receptors than ganglion cells, but the number of cortical neurons in area 17 is certainly orders of magnitude greater than the number of incoming fibres. The numbers at succeeding levels may be somewhat fewer, but a high proportion of the nerve cells in the brain must be capable of being influenced by vision, so if the hierarchical organization is pyramidal it is inverted rather than erect, divergent rather than convergent. If one uses the term 'cardinal cell', one must be sure to remember that the college of these cardinals outnumbers the church members and must include a substantial fraction of the 10^{10} cells of the human brain.

After-thoughts

It is sufficiently obvious that these propositions are incomplete, that there are aspects of the sensory problem left untouched, and that the dogmas go considerably beyond the evidence. I have said, in essence, that the cells of our brain are each capable of

more than had previously been supposed, and that what their activities represent may be more simply related to the elements of our conscious perceptions than had previously been thought. But clever neurons are not enough. The simplest computer program with its recursive routines and branch points has more subtlety than the simple hierarchy of clever neurons that I have here proposed as the substrate of perception.

I think one can actually point to the main element that is lacking. We have seen that some properties of the environment can be represented, or modelled, in a system of the type proposed; I feel that a corresponding model is also needed for our own motor actions and their consequences. Such motor and sensory models could then interact and play exploratory games with each other, providing an internal model for the attempts of our ever-inquisitive perceptions to grasp the world around us. A higher-level language than that of neuronal firing might be required to describe and conceptualize such games, but its elements would have to be reducible to, or constructible from, the interactions of neurons.

The five dogmas do not impede developments in this direction. My claim for them is that they are a simple set of hypotheses on an interesting topic, that they are compatible with currently known facts, and that, if any are disproved, then knowledge in this field will be substantially advanced.

References

Adrian, E. D., 1926a, "The impulses produced by sensory nerve-endings", Pt.1, *J. Physiol.*, **61**, 49–72.

Adrian, E. D., 1926b, "The impulses produced by sensory nerve-endings, Pt.4, Impulses from pain receptors", *J. Physiol.*, **62**, 33–51.

Adrian, E. D., 1928, *The Basis of Sensation* (Christophers, London; also Hafner, New York, 1964).

Adrian, E. D., 1941, "Afferent discharges to the cerebral cortex from peripheral sense organs", *J. Physiol.*, **100**, 159–191.

Adrian, E. D., 1947, *The physical background of perception* (Clarendon Press, Oxford).

Adrian, E. D., Zotterman, Y., 1926a, "The impulses produced by sensory nerve-endings, Pt.2, The response of a single end-organ", *J. Physiol.*, **61**, 151–171.

Adrian, E. D., Zotterman, Y., 1926b, "The impulses produced by sensory nerve-endings, Pt.3, Impulses set up by touch and pressure", *J. Physiol.*, **61**, 465–493.

Bard, P., 1938, "Studies on the cortical representation of somatic sensitivity", *Harvey Lectures 1938* (Academic Press, New York), pp.143–169.

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

Barlow, H. B., 1956, "Retinal noise and absolute threshold", *J. Opt. Soc. Amer.*, **46**, 634–639.

Barlow, H. B., 1960, "The coding of sensory messages" in *Current Problems in Animal Behaviour*, Eds. W. H. Thorpe, O. L. Zangwill (Cambridge University Press, Cambridge), pp.331–360.

Barlow, H. B., 1961, "Possible principles underlying the transformations of sensory messages" in *Sensory Communication*, Ed. W. A. Rosenblith (MIT Press, Mass. and John Wiley, New York), pp.217–234.

Barlow, H. B., 1969a, "Pattern recognition and the responses of sensory neurons", *Ann. N. Y. Acad. Sci.*, **156**, 872–881.

Barlow, H. B., 1969b, "Trigger features, adaptation, and economy of impulses", in *Information Processing in the Nervous System*, Ed. K. N. Leibovic (Springer-Verlag, New York), pp.209–226.

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

Barlow, H. B., Brindley, G. S., 1963, "Interocular transfer of movement after-effects during pressure binding of the stimulated eye", *Nature*, **200**, 1346–1347.

Barlow, H. B., FitzHugh, R., Kuffler, S. W., 1957, "Change of organization in the receptive fields of the cat's retina during dark adaptation", *J. Physiol.,* 137, 338–354.

Barlow, H. B., Hill, R. M., 1963, "Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects", *Nature,* 200, 1345–1347.

Barlow, H. B., Hill, R. M., Levick, W. R., 1964, "Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit", *J. Physiol.,* 173, 377–407.

Barlow, H. B., Levick, W. R., Yoon, M., 1971, "Responses to single quanta of light in retinal ganglion cells of the cat", *Vision Research,* 11, Suppl. 3, 87–102.

Barlow, H. B., Pettigrew, J. D., 1971, "Lack of specificity of neurones in the visual cortex of young kittens", *J. Physiol.,* 218, 98–100.

Bishop, P. O., 1970, "Beginning of form vision and binocular depth discrimination in cortex", in *The Neurosciences: Second Study Program,* Ed. F. O. Schmitt (Rockefeller University Press, New York), pp.471–485.

Blakemore, C., Campbell, F. W., 1969, "On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images", *J. Physiol.,* 203, 237–260.

Blakemore, C., Cooper, G. F., 1970, "Development of the brain depends on the visual environment", *Nature,* 228, 477–478.

Blakemore, C., Mitchell, D. E., 1973, "Environmental modification of the visual cortex and the neural basis of learning and memory", *Nature,* 241, 467–468.

Boole, G., 1854, *An Investigation of the Laws of Thought* (Dover Publications Reprint, New York).

Boring, E. G., 1942, *Sensation and perception in the history of experimental psychology* (Appleton Crofts, New York).

Burns, B., 1968, *The uncertain nervous system* (Edward Arnold, London).

Campbell, F. W., Green, D. G., 1965, "Optical and retinal factors affecting visual resolution", *J. Physiol.,* 181, 576–593.

Cragg, B. G., 1972, "The development of synapses in cat visual cortex", *Investigative Ophthalmology,* 11, 377–385.

Craik, K. J. W., 1943, *The Nature of Explanation* (Cambridge University Press, Cambridge).

De Valois, R. L., 1960, "Color vision mechanisms in the monkey", *J. Gen. Physiol.,* 45, Suppl., 115–128.

De Valois, R. L., Abramov, I., Jacobs, G. H., 1966, "Analysis of response patterns of LGN cells", *J. Opt. Soc. Am.,* 56, 966, 977.

De Valois, R. L., Abramov, I., Mead, W. R., 1967, "Single cell analysis of wavelength discrimination at the lateral geniculate nucleus in the macaque", *J. Neurophysiol.,* 30, 415–433.

Dreyfus, H. L., 1972, *What Computers Can't Do* (Harper and Row, New York).

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

Ewert, J. P., 1970, "Neural mechanisms of prey-catching and avoidance behavior in the toad (*Bufo bufo* L.)", *Brain Behav. Evol.,* 3, 36–56.

Fry, G. A., 1948, "Mechanisms subserving simultaneous brightness contrast", *Am. J. Optom.,* 25, 162–178.

Gabor, D., Kock, W. E., Stroke, G. W., 1971, "Holography", *Science,* 173, 11–23.

Galambos, R., 1944, "Inhibition of activity in single auditory nerve fibers by acoustic stimulation", *J. Neurophysiol.,* 7, 287–303.

Galambos, R., Davis, H., 1943, "The response of single auditory-nerve fibres to acoustic stimulation", *J. Neurophysiol.,* 7, 287–303.

Galambos, R., Davis, H., 1948, "Action potentials from single auditory-nerve fibres?" *Science,* 108, 513.

Ganchrow, J. R., Erickson, R. P., 1970, "Neural correlates of gustatory intensity and quality", *J. Neurophysiol.,* 33, 768–783.

Granit, R., 1947, *The Sensory Mechanisms of the Retina* (Oxford University Press, Oxford).

Granit, R., Svaetichin, G., 1939, "Principles and technique of the electrophysiological analysis of colour reception with the aid of microelectrodes", Upsala Läkaref Färh., 65, 161–177.

Green, D. M., Swets, J. A., 1966, *Signal Detection Theory and Psychophysics* (John Wiley, New York).

Gross, C. G., Rocha-Miranda, C. E., Bender, D. B., 1972, "Visual properties of neurons in inferotemporal cortex of the macaque", *J. Neurophysiol.,* 35, 96–111.

Grüsser, O.-J., Grüsser-Cornehls, U., 1968, "Neurophysiologische Grundlagen visueller angeborener Auslösemechanismen beim Frosch", *Zeitschrift für vergleichende Physiologie,* 59, 1–24.

Hartline, H. K., 1938, "The response of single optic nerve fibres of the vertebrate eye to illumination of the retina", *Am. J. Physiol.,* 121, 400–415.

Hartline, H. K., 1940a, "The receptive fields of optic nerve fibers", *Am. J. Physiol.,* **130,** 690-699.

Hartline, H. K., 1940b, "The effects of spatial summation in the retina on the excitation of the fibers of the optic nerve", *Am. J. Physiol.,* **130,** 700-711.

Hecht, S., Shlaer, S., Pirenne, M., 1942, "Energy, quanta, and vision", *J. Gen. Physiol.,* **25,** 819-840.

Hirsch, H. V. B., Spinelli, D. N., 1970, "Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats", *Science,* **168,** 869-871.

Hirsch, H. V. B., Spinelli, D. N., 1971, "Modification of the distribution of receptive field orientation in cats by selective visual exposure during development", *Exp. Brain Res.,* **13,** 509-527.

Hubel, D. H., Wiesel,T. N., 1959, "Receptive fields of single neurones in the cat's striate cortex", *J. Physiol.,* **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.,* **160,** 106-154.

Hubel, D. H., Wiesel, T. N., 1963, "Receptive fields of cells in striate cortex of very young, visually inexperienced kittens", *J. Neurophysiol.,* **26,** 994-1002.

Hubel, D. H., Wiesel, T. N., 1965, "Binocular interaction in striate cortex of kittens reared with artificial squint", *J. Neurophysiol.,* **28,** 1041-1059.

Hubel, D. H., Wiesel, T. N., 1970, "The period of susceptibility to the physiological effects of unilateral eye closure in kittens", *J. Physiol.,* **906,** 419-436.

Hurvich, L. M., Jameson, D., 1960, "Perceived color, induction effects, and opponent-response mechanisms", *J. Gen. Physiol.,* **43,** Suppl., 66-80.

Ingle, D., 1968, "Visual release of prey-catching behaviour in frogs and toads", *Brain, Behaviour and Evolution,* **1,** 500-518.

Ingle, D., 1971, "Prey-catching behaviour of anurans toward moving and stationary objects", *Vision Research,* Suppl. No.3, 447-456.

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

Lashley, K. S., 1929, *Brain Mechanisms and Intelligence: a Quantitative Study of Injuries to the Brain* (University of Chicago Press, Chicago).

Lashley, K. S., 1950, "In search of the Engram physiological mechanisms in animal behaviour", in *Symposium of the Society for Experimental Biology,* Ed. J. F. Danielli and R. Brown (Cambridge University Press, Cambridge).

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. Rad. Eng.,* **47,** 1940-1951.

Levick, W. R., 1967, "Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit's retina", *J. Physiol.,* **188,** 285-307.

Mach, E., 1886, *The Analysis of Sensations, and the Relation of the Physical to the Psychical* Translation of first edition (1886) revised from fifth German edition by S. Waterlow (Open Court, Chicago and London, 1914) Ed. C. M. Williams. (Also Dover Publications, New York, 1959.)

Marshall, W. H., Woolsey, C. N., Bard, P., 1941, "Observations on cortical somatic sensory mechanisms of cat and monkey", *J. Neurophysiol.,* **4,** 1-24.

Maturana, H. R., Lettvin, J. Y., McCulloch, W. S., Pitts, W. H., 1960, "Anatomy and physiology of vision in the frog (*Rana Pipiens*)", *J. Gen. Physiol.,* **43,** Suppl. No.2, *Mechanisms of Vision,* 129-171.

McCulloch, W. S., 1959, "Agatha Tyche: of nervous nets—the lucky reckoners", in *Mechanisation of Thought Processes: Proceedings of a Symposium Held at the National Physical Laboratory, Vol.2* (HMSO, London), pp.611-634.

Pettigrew, J. D. (forthcoming), "The effect of visual experience on the development of stimulus specificity by kitten cortical neurons".

Pettigrew, J. D., Freeman, R. (forthcoming), "Visual experience without lines: Effect on developing cortical neurones".

Pettigrew, J. D., Nikara, T., Bishop, P. O., 1968, "Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence", *Exp. Brain Res.,* **6,** 391-410.

Pfaffman, C., 1941, "Gustatory afferent impulses", *J. Cell. Comp. Physiol.,* **17,** 243-258.

Pfaffman, C., 1955, "Gustatory nerve impulses in rat, cat, and rabbit", *J. Neurophysiol.,* **18,** 429-440.

Sakitt, B., 1972, "Counting every quantum", *J. Physiol.,* **222,** 131-150.

Shannon, C. E., Weaver, W., 1949, *The mathematical theory of communication* (University of Illinois Press, Urbana).

Sherrington, C. S., 1941, *Man on His Nature* (Cambridge University Press, Cambridge).

Shlaer, R., 1971, "Shift in binocular disparity causes compensatory change in the cortical structure of kittens", *Science,* **173**, 638–641.

Stone, J., 1972, "Morphology and physiology of the geniculocortical synapse in the cat: The question of parallel input to the striate cortex", *Invest. Ophthal.,* **11**, 338–346.

Svaetichin, G., MacNichol, E. R., Jr., 1958, "Retinal mechanisms for chromatic and achromatic vision", *Ann. N. Y. Acad. Sci.,* **74**, 385–404.

Talbot, W. H., Darian-Smith, I., Kornhuber, H. H., Mountcastle, V. B., 1968, "The sense of flutter-vibration: Comparison of human capacity with response patterns of mechano-receptive afferents from the monkey hand", *J. Neurophysiol.,* **31**, 301–334.

Tanner, W. P., Jr., Swets, J. A., 1954, "A decision making theory of visual detection", *Psychol. Review,* **61**, 401–409.

Wagner, H. G., MacNichol, E. R., Wolbarsht, M. L., 1960, "The response properties of single ganglion cells in the goldfish retina", *J. Gen. Physiol.,* **43**, Suppl., 115–128.

Wiesel, T. N., Hubel, D. H., 1963, "Single cell responses in striate cortex of kittens deprived of vision in one eye", *J. Neurophysiol.,* **26**, 1004–1017.

Wiesel, T. N., Hubel, D. H., 1965, "Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens", *J. Neurophysiol.,* **28**, 1029–1040.

Wohlgemuth, A., 1911, "On the after-effect of seen movement", *Brit. J. Psychol.,* Monograph. Suppl., **1**, 1–17.

Woodward, P. M., 1953, *Probability and Information Theory with Applications to Radar* (Pergamon Press, Oxford).

Yarbus, A. L., 1965, *Eye Movements and Vision,* Translated from Russian by Basil Haigh (Plenum Press, New York).

Zangwill, O. L., 1961, "Lashley's concept of cerebral mass action", in *Current Problems in Animal Behaviour,* Eds. W. H. Thorpe, O. L. Zangwill (Cambridge University Press, Cambridge).