7 Modes of perception

Perception, very broadly, refers to the utilisation of the organs of sense. Human perception, though, is often defined in terms of the use of the senses to acquire conscious knowledge of the outside world, with the existence of a human mind being assumed as the destination which sense data may reach. Less tendentiously, we could discuss perception, in a more modern phrase, as the processing of information received via the sense organs. Even in this case, there are many problems in trying to compare the perception of men and animals. For human perception, the inner organisation of knowledge needs always to be considered as one of the determinants of what is sensed. Context, expectancies, the synthesis of pieces of sensory information into coherent perceptual wholes: these are variables of crucial importance in theories of human information processing—but to what extent do these things affect the perceptions of animals? At the extreme, human perception may be characterised as a sequence of intentional perceptual acts, requiring selection, choice and effort. Does this mean that human perception is fundamentally different from the activation of the sense organs of animals? Philosophical views on fundamental differences between human and animal perception, particularly those of Locke, have been discussed in Chapters 1 and 2. Here I shall attempt to start again from the beginning: using first, information from laboratory tests of the perceptual capacities of animals, by which some of the questions about the limits of these capacities can be answered by reference to behavioural data; and second, modern knowledge, such as it is, of how brain processes are involved in perception.

Modalities and qualities in perception

The easiest distinctions to make about perception concern the method of operation of the external sense-organs. The eyes are clearly different from the ears and we can say that acquiring information through the eyes is vision, and acquiring information through the ears is hearing, without fear of contradiction. This is a distinction between sensory modalities, and it is obvious that we can separate out five modalities of external perception—vision, hearing, touch, taste and smell—according to whether the eyes, the ears, the skin, the tongue or the
nose can be identified as the source of particular sensations.

This is simple enough, but once we begin to enquire as to what kinds of information can be transmitted by each modality, things become rather more complicated and subjective. These types of information can be discussed as stimulus qualities, and it should be emphasised that stimulus qualities which are subjectively obvious to ourselves can only be attributed to the operations of perception in other species on the basis of physiological and behavioural evidence. One stimulus quality is common to all modalities and that is intensity. We should expect sense organs to be capable of reacting differentially to a bright light or a dim light, a loud or soft sound, intense or mild pressure on the skin, and so on. As far as the transformation of external events by sense organs goes, we may safely assume that differentiation between bright and dim lights is not a uniquely human accomplishment, although whether a bright light is perceived as having something in common with a loud noise in any other species is something which would require experimental confirmation.

There are many stimulus qualities which are specific to particular modalities, and there are some, like intensity, which may or may not apply to more than one modality, depending on the capabilities of the animal involved, and perhaps on prior experience. Colour, and certain aspects of brightness and shape, apply to vision; pitch and timbre to hearing; sweetness and bitterness to taste. Interactions between these are an interesting aspect of human subjective experience—we speak of warm colours and sweet sounds—and have been investigated under the heading of ‘synaesthesia’, but there are little or no data on whether similar subtleties might be apparent to animals. However, the independence of the mental analysis of stimulus qualities from the immediate activities of particular sense organs is clearly one of the ways in which ‘higher’ forms of perception may differ from ‘lower’. This sort of distinction is one I wish to emphasise by referring to different ‘modes’ of perception.

Reflexive and cognitive modes of perception

The use of the term ‘modes’ as well as ‘modalities’ may appear unnecessarily confusing, but the distinction I want to make is really a very straightforward one. It is possible to purchase equipment which will arrange that one’s garage door opens whenever a car appears before it—similar devices make walking through the compartments of trains very much easier than it used to be. On simple behavioural grounds it might seem that the doors, or the sensing devices connected to them, are ‘perceiving’ the cars or persons which present themselves in appropriate places. No one would suppose, however, that such perception has a great deal in common with the processes which take place in our own brains. It might be claimed, in the Cartesian tradition, that such mechanical or electronic arrangements should be taken as analogies for animal, as distinct from human, perception. It is necessary, therefore, to attempt some sort of categorisation of modes of perception that do
or do not mirror various aspects of our own perceptual abilities.

**Reflexive perception**

To start with, we can pick out the simplest possible kinds of reactions to stimuli, reactions analogous to the opening of a door in response to the breaking of a photocell beam. It is unlikely that any response of organic tissue is quite as simple as this, but, clearly, some reactions of very simple animals to external stimulation, and some of the reflexive responses of more complicated animals, ought to be separated out. A single cell, either as a single-celled animal, or as a specialised part of a mammal, may respond to touch, temperature, or chemical stimulation, in particular ways. The spinal cord of vertebrates, even in isolation from the brain, allows for a variety of limb movements in response to touch, and for reactions of the digestive system to internal forms of stimulation. There is a sensitivity to events here, which is a characteristic of perception; but the nature of the stimulating events, and the individual response to them, is relatively fixed and rigid. The fixity and rigidity is relative to the greater complexity and malleability of higher forms of perception—it would be a mistake to oversimplify the characteristics of even these simpler forms of organic response to stimulation. We can assess the reflexiveness of these stimulus-response relationships both in terms of the complexity of the stimulating event, and of the flexibility of the forms of response to it. In some cases the information which is ‘recognised’ even by a single cell may come in quite an elaborate form, as in the ‘recognition’ by mammalian immune mechanisms of cells which belong to the individual animal rather than to another animal of the same species. In other cases, as in the scratching reflexes of decerebrate frogs, the perceptual input may appear rather rudimentary, but the nature of the response to the input may be complex. The way in which tactile end-organs pick up chemical or mechanical stimulation at some point on the skin of a leg of the decerebrate frog is easily explained: the way in which the spinal cord reacts by bringing up the other foot to just this point to make wiping movements has to be interpreted in terms of quite elaborate motor organisation.

But in all these cases there is a restricted category of stimulating event and the only way in which this stimulus can be said to be perceived is that there is an immediate and pre-programmed response to it. It is improper to think of even spinal reflexes as completely fixed and isolated from one another, but there is an important distinction to be made between, on the one hand, the sensing of stimulation which takes place in the organised movements of swallowing, the unconscious co-ordination of particular muscles in walking or standing, or the accommodation of the lens of the eye to objects at different distances, and, on the other hand, such things as the perceptual experience of recognising another person in a photograph or noticing that traffic lights have
Cognitive perception

The characteristics of human perception of events include some kind of detached knowledge which is not tied exclusively to a single form of response, and which is typically (though not universally) available to some system of verbal description. Perception in this sense is closely linked to memory—we not only perceive that the traffic lights have changed, but remember that we have just perceived this; we not only recognise someone, but we know *that* we have recognised him. This is not to deny the reality of unconscious or ‘subliminal’ forms of human perception, but to acknowledge that human perception usually needs to be accounted for in terms of complex inner mental organisation. The more serious theories of human perception have to say that when a person sees something, he acquires a belief, or has an idea, about what it is that he is seeing. Such a concept is made more technical, if it is said

that visual information is assimilated into a ‘schema’ (the term used by Piaget and Neisser), but the main thing is that there is a cognitive aspect of perception and this is quite different from reflexive reactions to stimulation such as the contraction of the pupil in response to bright lights or a jerk of the arm when the finger touches a burning-hot plate.

Modes, qualities and modalities in human and animal perception

There will surely be little disagreement that distinctions can be drawn between reactions to events of a ‘knee-jerk’ reflex type, and more knowledgeable ways of extracting information from the environment by the use of the senses. There is considerable difficulty, however, in coming to decisions about whether the utilisation of the senses by animals falls into one category or the other. Part of the difficulty arises, I believe, from the tradition, examined in earlier chapters, of assuming that all animal perception is of the reflex type unless it can be shown otherwise. That this assumption is unrealistic and misleading can be shown by examining the functional relationship between the modality of sensory input and the perception of stimulus qualities.

The essence of reflexive perception is that information is confined within the modality in which it is received. A touch on the skin is a touch on the skin, rather than the presence of another animal. But is this the way that the senses are normally utilised by, for instance, the cat? If my cat is sitting down, and I lightly snap my fingers behind her head, I may observe one ear, and then both ears, turning towards the sound. This can be put down to a sophisticated auditory reflex. If I snap my fingers a little louder, however, I would expect to see not just the ears, but the whole head turn with the eyes firmly fixed on my hand. And it would not be surprising if the whole animal then turned and moved, and came and sniffed the hand and rubbed her head against it, making contact with the
whiskers just before the final rub with the cheek. Would it be reasonable to construe this sequence as a succession of reflexes, each isolated in particular modalities of hearing, vision, smell and touch? I suggest not, and that on the contrary, the brain of the cat is organised so as to connect and correlate the various modalities, one with another.

The information within one modality may take part in various specialised reflexes, but can also be put towards an interpretation, schema, or belief about the current state of affairs in the world which is cognitive because it can be applied across modalities and across a certain range of changes of time and place. After rubbing one cheek against my hand, the cat may turn around, losing sight of the hand, and come back to rub its other cheek. In ordinary language, the cat knows that the hand is there. I am not suggesting that the cat knows as much about me as I know about the cat, but I would argue that, in terms of immediate perception, the cat knows that the hand is there in much the same way that I know that the cat is there.

The most abstract and overriding, and the most useful, perceptual quality which is independent of modality is something that may be termed object identity. By this I mean that the cat perceives a ‘thing’ rather than a sensation, and that its perception of the thing must consist of a collection of bits of knowledge. The hand is something that looks solid, can be rubbed against, will smell and feel like a hand, and will do the things hands normally do. This is not to claim that the cat mentally runs through a list of this type every time it sees a hand, but that when it sees a hand, a complex mental process of recognition takes place. This sort of hypothesis may be theoretically debatable, and difficult to subject to experimental testing. But I believe that we know enough about the behaviour of mammals, and the brain mechanisms available to them, to make it extremely plausible, and that experimental data of the kind to be discussed below will support it.

If the extent to which animals analyse (and synthesise) their environment into ‘things’ is difficult to determine, it may be easier to isolate rather less abstract stimulus qualities which, while they may serve as attributes of things, can be derived more obviously from correlations between sensory modalities, or from specific features within sensory modalities. Isolating sensory qualities is certainly a more speculative matter than differentiating between modalities, and no doubt there are considerable differences between species or between classes of animal, but let us look at the possibility that animals may perceive such qualities as location, movement and object identity with a certain degree of independence from particular modalities.

Location and movement

Location presupposes some minimum case of object identity, since one has to ask ‘location of what?’ But given some rudimentary
classification of stimulus events such as the identification of a touch on the skin, or the identification of an insect via the visual ‘bug detectors’ of a frog, then these events may be given a location in terms of ‘where exactly is the body surface touched?’ or ‘where is the visual field and at what distance is the detected bug?’ Now in the case of the scratching reflexes of a frog, or the tongue-throwing response of the same animal, it is possible that location is not a cross-modal quality, since receptors on the skin surface may have rather direct connections with the muscles used in the scratching reflex, and the transformation of data from the eye which indicate the angle and distance for effective tongue-throwing may take place in a reflexive and automatic way. But in the case of mammalian predators locating prey, then scenting, sighting, hearing, and touching might be expected to work in a co-operative fashion. This is especially obvious if the activation of one modality leads to orientation, or investigation with the others. The co-operation of the eyes and ears of the cat has already been emphasised, but a correlation between visually detected closeness and touch with the whiskers and skin seems equally likely. In mammals with weak vision, which are often classed as ‘primitive’, such as the hedgehog, or the mole, the joint use of smell and touch to establish ‘cognitive maps’ may be more important than the combination of vision and hearing.

Either for touch, or for vision, perception of the position and posture of the animal’s own body has to be taken into account for the accurate perception of the external world; If we look at a door while turning our head to one side, we do not sense the door turning, although the image on our retina moves: the inner sensing of the movement of our head is used to compensate for the movement of the retinal image. If I close my eyes and touch several mugs on the table, I know which mug is closest because it is the mug I touched when my arm was flexed, even though the tactile sensation in my fingers might be the same in this case as in touching a mug with my arms stretched out. It is highly unlikely that other mammals manage without a very similar kind of co-operation and compensation between different forms of perception, and the assessment of external information in terms of the position and movement of the body of the receiving animal is a necessity throughout most of the animal kingdom.

The interaction between internal and external senses is especially necessary in the perception of external movement. Specialised receptors respond when a visual array moves across the retina, or a tactile stimulus moves past the whiskers or skin. But animals need to know, surely, whether it is they or external objects, that are moving.

Incoming information from external-movement detectors must be constantly monitored in terms of the motor activities currently being performed. This need not count as a particularly cognitive form of perception—it is something taken for granted as a continuous and
unconscious adjustment necessary to maintain human perceptual reality (otherwise the room would seem to spin every time we turned our heads). However, it can be used to illustrate the fact that very basic ways of perceiving and interacting with the environment are not as simple as they seem, and that the perceptual apparatus utilised by vertebrates simply for movement from place to place may require a considerable degree of internal organisation and interpretation.

**Object identity and value**

Neisser (1976) has stressed the analogy between human perception and the activity of picking apples from a tree. Apples do not attempt to force themselves down our throats, he says, and neither do perceptual experiences force themselves upon us. Rather, we ourselves select and choose among sense impressions. On what basis do we make the choice, and to what extent can animal perception be said to be similarly a matter of selection? One of Neisser’s examples was the selective ability of human subjects watching sporting events on a television screen to attend to a coherent sequence of visual images, even in the face of confusion introduced by the superimposition of recordings of more than one game on the same screen. Overtly passive watching of a television screen can be taken as a supremely ‘cognitive’ mode of perception, in so far as information may be absorbed in the absence of any direct effect on behaviour, reflexive or otherwise. But the input of information is clearly not emotionally neutral—one could say that even the minimal effort required to maintain eye-contact with the screen is maintained only so long as there is at least a momentary interest in the scene depicted and that passionate involvement with unfolding human dramas thus represented is not unusual.

Assigning values to purely visual images may be a peculiarly human characteristic—the question is not whether animals could show an equivalent interest in moving pictures, but whether there is an element of choice and direction in the way they gather information from the real environment. Do animals categorise certain perceptions as good or bad, desirable or undesirable, to be sought out or to be avoided? It would seem hard to deny that there are motivational or emotional connotations to perceptions that are related to basic drives of hunger, thirst and reproduction but there are clear differences of opinion as to whether the emotional aspects of such perceptions work in a reflexive or cognitive way. A hungry animal must be said to perceive food, if it eats the food, but it is often difficult to decide whether, for instance, an innate reaction to the smell of food leads to reflexes of biting, salivating and swallowing, coupled with a degree of temporary excitement and emotional arousal, or whether, as well as this, there is a more dispassionate perception of food as a ‘thing which may be eaten’.
This was one of the questions addressed in Chapter 3, in the comparison between stimulus-response and cognitive theories of animal behaviour, where I argued that laboratory rats and pigeons, to say nothing of monkeys and chimpanzees, are capable of remembering and anticipating both perceptions of food, and such things as where the food should be, and what actions might be necessary to obtain it. If this argument is accepted, the value of objects is clearly one of the determinants of object perception by animals. At its simplest, this is an assertion that animals categorise objects into ‘good to eat’ or ‘not good to eat’: by reviewing laboratory experiments on animal perception one might be led to believe that this is the only perceptual categorisation of objects ever made by sub-human species. It may be the case that a large part of animal perception is constrained by the Darwinian imperatives of searches for food, safety and social experience, but these constraints are undoubtedly overemphasised by the more abstruse restrictions of the search by animal psychologists for effective experimental techniques. in particular, there is very much more laboratory data on what hungry animals can perceive about stimuli associated with experimental food rewards than there is on what animals normally perceive in the course of their natural movements and social interactions. Another limitation of the available experimental data is that it is heavily weighted towards vision. This is partly because it is technically much easier to present animals with particular visual arrays than it is to arrange controlled delivery of specific smells, tastes or tactile sensations, but also because the modality of vision involves especially interesting theoretical problems. In a subsequent section I will review a number of experimental findings which bear on the degree of cognitive organisation involved in visual perception by animals.

Cognitive organisation in animal perception

The application of rigorous experimental techniques to the study of animal perception has had rather mixed effects. On the one hand, unequivocal evidence is available to document the sensitivity and functional effectiveness of animal sensory systems. In some cases, such as Kalmus’s careful study of the ability of dogs to identify human individuals by scent (Kalmus, 1955) the results reaffirm conclusions arrived at by more naturalistic observations. In others, such as the experiments in Pavlov’s laboratory (discussed on pp. 65W) which established the sensitivity of dogs to the rhythm, timbre and sequence of musical notes, systematic investigation has allowed the detection of sensory capacities which might have gone unnoticed by the more casual eye. There is no doubt about the advantage of conventional and reliable experimental techniques in instances such as these. But if there has been a drawback to the use of standard experimental procedures, it has been in the way in which the restricted performances demanded of experimental animals have given apparent support to the absurd theoretical proposition that the only variation possible in an animal’s reception of external events was in the instigation, or non-instigation, of an immediate overt behavioural reaction.
In Pavlov’s experiments on the conditioning of salivation to various sounds, for instance, one dog might salivate to the note of middle C whether played on a clarinet, organ pipe, or tuning fork, but not to any note more than a semi-tone different in pitch, however played, because only middle C signalled food. But another dog might salivate to any note whatever played on a clarinet, and not to any sound made by an organ pipe or tuning fork, because for this dog, notes from a clarinet but not from any other instrument had been used as the indication for food. Pavlov himself spoke of such results in terms of lower and higher kinds of auditory analysis. Subsequently, however, there has been a tendency for other theorists to ignore such questions, and to make the assumption that the perception of the sound stimuli must occur in an identical way for both dogs (and that in so far as it is mentioned, the dogs listen and hear in the same way), the difference between the dogs salivating to different sounds being that the response of salivation attaches itself to a different subset of possible auditory patterns (e.g. Spence, 1936). It may not seem as though the form of this description’s very important, but it means for one thing that the theoretical emphasis very much concerns the response which happens to be used in a particular experiment; the problem with this is that the overt response (in this case salivation) may be the least interesting aspect of the perceptual ability being studied (in this case hearing). More specifically, the response-based description misses out the possibilities of perceptual learning, and selection and choice within perception. Another form of description would emphasise that the first dog had learned to pay attention to pitch, while the second had learned to recognise the sound of a clarinet, but chose, for good reason, to ignore pitch.

In fact, there is overwhelming evidence suggesting that the second form of description, that which emphasises selection and learning during the process of perception, is more accurate than the first, which is limited to responses which happen to be made as a consequence of stimulus input. Species commonly used in the laboratory, such as dogs and cats, rats and pigeons, can be shown to organise and direct their perceptual activities independently of the movements and bodily reactions measured to assess these. The evidence is extensively reviewed by Sutherland and Mackintosh (1971) and Mackintosh (1974), and can be put under two general headings. In the first place, the effects of perceptual training are not inextricably tied to particular overt responses. Suppose the two hypothetical dogs described above, one trained to salivate to middle C whatever instrument it is played on, and the other trained to salivate to the sound of a clarinet irrespective of pitch, were given a new test, which did not involve salivation, but which required the same perceptual discriminations. This is easier said than done, but we may imagine the two dogs lifting their paws in response to a middle C, or to a clarinet note, respectively. We should definitely expect some transfer of the effects of perceptual training, so that the animal which was already in the habit of noticing middle C should find it easy to apply this perceptual habit to a new task, and the dog
which had already learned to recognise the clarinet should continue to recognise it even when salivation was not called for. Comparisons with animals lacking the previous training would of course be needed, but I think few theorists would really want to stake very much on the claim that recognition of a clarinet takes place in the salivary gland, when it came to the point.

Actual experiments along these lines are rare, but Lawrence (1949, 1950) successfully demonstrated transfer of perceptual learning between different response tasks using visual discriminations in rats. The fact that vision in rats is very poor, and is not particularly amenable to

the effects of experimental training gives these results additional weight. A much more frequently performed type of experiment is one in which there is a change in response output with similar perceptual requirements, simply reversing the response rule. The dog trained to salivate to the sound of a clarinet, but not to an organ pipe, could be retrained so that the organ pipe was the positive stimulus, leading to food, and the clarinet was the negative one, no longer signalling the food. The usual initial effect of such a reversal is total confusion, although oddly enough there are certain conditions under which rats with long experience of a difficult visual discrimination can accomplish a relatively rapid reversal of response (the ‘overtraining reversal effect’, see Mackintosh, 1974, pp. 602—7). However, a standard kind of experiment, which gives exceedingly reliable results, is to wait until responding has settled down after the first turn about, then to reinstate the original conditions, then to reverse the stimuli again, and so on. The results of this ‘serial reversal learning’ procedure are that the long-suffering animals eventually become adept at switching into the required pattern of responding at a moment’s notice, making few mistakes at any stage. Chickens and pigeons, for instance, will learn to peck a red disc but not a green one as long as red is ‘correct’, and indicates that food is available; and then to switch to pecking at green, but not red, as soon as the experimenter reverses the conditions (Levine, 1974). There are various other implications of this, to do with the ability to remember which signal is correct at any particular moment (see Mackintosh, 1974, pp. 608—10); but clearly the animals must pay close attention to the colour of the stimuli in the absence of a consistent code by which a single colour is connected to a single response.

Apart form directly chopping and changing between the responses animals are supposed to make as a consequence of their perceptions, the independence of perceiving and responding can be illustrated by the transfer of perceptual effects from one set of stimuli to another. It can be argued, for instance, that training a dog to salivate only to middle C must have had the general effect of inducing attention to pitch, if subsequent training to salivate to other individual notes is facilitated. Also, we can certainly say in general that it is possible to draw an animal’s attention to a particular sensory modality by the use of rewards or punishments—pigeons which show no sign of noticing tones played into a Skinner box
where they are performing a learned response rapidly acquire a
sensitivity to the auditory quality of the

tones if the presence of a tone is made significant as a predictor of
their food rewards (Jenkins and Harrison, 1960). Much more
detailed focusing of attention is possible within particular
modalities, however, as in the case of selective attention to the pitch,
or intensity, of sounds; or the shape, instead of the colour, of visual
displays. Given some variety of visual stimuli in initial training,
such as presentations of squares and circles which may be red or
yellow, then animals trained to pick out red, but not yellow displays,
irrespective of shape, are, not surprisingly, more ready to make
distinctions between other colours (blue and green, say) than
animals set up to ignore the original colours, and pick out individual
shapes (see Mackintosh, 1974, pp. 597—8, on ‘Intradimensional
and extradimensional shifts’). Thus animals may be trained not only
to pay attention to one modality rather than another, but also to
notice (or ignore) certain qualities within a particular modality.

Various experimental permutations and combinations of the
kinds of stimuli presented to animals, and the nature of the overt
responses the animals are persuaded to make when the stimuli
occur, all support the hypothesis that the perception of the stimuli is
in some sense independent of particular overt responses made to the
stimuli. In other words the mode of perception demonstrated in
these experiments is cognitive rather than reflexive, using the terms
as discussed above. The degree to which it is cognitive may be
rather limited, but it is well worth stressing that animals do not just
make certain responses to immediately present events, but
selectively and actively perceive things about stimuli, which allow
them to behave appropriately under experimental conditions. In
Piaget’s terms, stimuli can be ‘accommodated’, leading to the
formation of new perceptual schemata, rather than merely
‘assimilated’ into currently available reactions.

Perhaps the clearest demonstration of the detached and selective
character of animal perception, in contrast to the notion that
environmental events force themselves willy nilly through the
sensory receiving apparatus and out to automatic responses, is the
finding that experimental subjects can change, as it were at will,
from noticing one thing to noticing another about the same set of
sensory information before them. One kind of procedure which
allows this sort of demonstration is referred to as ‘conditional
discrimination’ and an experiment reported by Reynolds (1961)
serves as an example. Pigeons were shown four visual stimuli, in a
repeated succession. These stimuli consisted of white shapes in a
coloured background: a blue
circle, a blue triangle, a red circle and a red triangle. One way to
demonstrate the perceptual combining of visual features would
simply be to train the birds to respond to only one of these stimuli—
the blue triangle say—which could be done with no difficulty. What
Reynolds did instead was to show the perceptual isolation of the
features of colour and shape in the visual displays in front of the birds, by training them, at a signal, either to select both triangles, irrespective of colour (by pecking at them), or to select both red stimuli, irrespective of shape. The instruction as to which selection to make was given by the illumination of different light bulbs to the side of pigeons. Logically, performing this task might seem to be an unnecessarily complex achievement for pigeons as a species to have evolved, or for individual animals to trouble to demonstrate. Functionally, however, it makes sense for visual analysis to proceed on an if-then basis, to the extent of for instance putting the detection of hawk-like features along with things moving in the sky, but grain-like features together with little things lying on the ground; and, more generally, the use of visual perception for such activities as foraging for food may be advanced by selective capabilities. It has often been suggested that foraging birds and mammals may activate ‘search images’, so that they may be looking for moths at one point, and beetles at another (Croze, 1970; Murton, 1971; Krebs and Davies, 1978). This means moths will be more readily detected while they are being searched for, even though other prey are then more likely to be missed. If this is so, then in terms of perceptual organisation it means that there must be a moth-schema, or a moth-recognition procedure, that can be selectively brought into play to modify from within receipt of messages from without.

*Dimensions, analysers, descriptions and representations*

Even the bare facts of the visual recognition of objects at different distances and in different places in the visual field require a theory of perceptual analysis of some sophistication (Sutherland, 1959). This is not quite so clear in the case of other modalities; but if we rub our palms over the corner of a table, and then feel the same corner with our fingers or forearm, there is a ‘corner’ quality to the sensation which is generalised over the body surface, and over different intensities of pressure, which corresponds to ‘pattern recognition’ at different locations and distances in the visual field; and we expect a dog to recognise ‘his master’s voice’, or the sound of a clarinet, as a category of auditory pattern, despite considerable variations in the dimensions of absolute pitch, and intensity.

The general type of theory put forward to account for the facts of perceptual recognition, especially in the case of vision, is known as ‘feature analysis’ (Selfridge, 1959; Sutherland, 1959, 1968; Frisby, 1979). Such a theory presupposes the existence of a hierarchy of analysers, each level capable of detecting perceptual features progressively further removed from the detailed reality of neural activity transmitted from the sense organs. The flow traditional example is the invariant recognition of the many different patterns of light impinging on the retina of the human eye which may be classified as a letter ‘A’ (e.g. Neisser, 1967). At the first
level of analysis, we might imagine detectors of light and shade at every point in the visual field. Then we would need a level at which lines were distinguished from background, and the length and angle of dark or light lines on the retina coded. After this, we must have detectors for various arrangements of lines, which can code such aspects of the arrangements as ‘two legs at the bottom’, ‘horizontal cross bar in the middle’, ‘corner point at the top’. Finally we end up with an ‘A’ analyser, which recognises ‘A’s of various sizes, and in different typefaces and handwritings. This last stage is by far the most perplexing, since it must put in the same category capital ‘A’s with round or pointed tops, printed lower case ‘A’s, and the enormous range of handwritten capital and lower case ‘A’s. In practice, and especially for difficult-to-read handwriting, we should expect our ‘A’ analyser also to guess the presence of ‘A’s on the basis of context, and surrounding information, as in the sequence T-H- squiggle-T. This means that a single analyser has been given rather a big job, and it is arguable that in saying that there exists a final ‘A’-analyser, we are simply translating the fact that we can recognise ‘A’s into a different terminology.

Sutherland (1968) broadened the concept of the final analyser by referring to ‘structural descriptions’ which enable a comparison to be made between current patterns of stimulation and previously established criteria of form. As with final analysers, the crucial point about structural descriptions is that they include several quite different alternatives and incorporate wide ranges of variation. The internal structural description of an ‘A’ would have to contain a similar amount of information as is given in the form ‘Anything either with two angled lines meeting in a point at the top with a cross bar half way down or something similar with a rounded top or a circle with a vertical line of the same height touching at the right or etc’. Clearly one does not expect such a description to be written out in code somewhere in the brain, but the structural aspect of the description is that neurons or sets of neurons in the brain correspond in function to analysers or feature detectors, at various levels of the hierarchy of analysis, and are connected together in parallel in such a way that the possibilities of a serial, written out, description are mimicked. The emphasis on structure as a fixed set of relationships is in some ways unfortunate, however, since one of the peculiarities of perception which I have just emphasised in the case of pigeons is that descriptions are flexible, and can be rapidly changed. The emphasis in Sutherland and Mackintosh’s account of analysers as they are theoretically derived from experimental studies of animal perception (Sutherland and Mackintosh, 1971) is that analysers are ‘switched in’ or ‘switched out’ according to their current usefulness to the animal. In human perception, direct instructions such as ‘look for all the vowels, and ignore capital letters’ radically alter the way in which visual information is processed.

I shall prefer, therefore, to use the term ‘inner description’ to refer to the criteria which, as an aspect of perceptual processing in the brains of animals, are used as the basis for recognition and
classification of experienced events. Another rather similar theoretical term is ‘representation’ often used to imply that an animal is assumed to conjure up some sort of remembered perception of an event (e.g. Rescorla, 1979, 1980). It will be convenient to use these two terms, ‘inner description’ and ‘representation’, to make the kind of distinction emphasised by Bindra (1976, see pp. 99—102, this volume) between relatively permanent cognitive organisation and an activated memory, or image. ‘Inner description’ is my preferred version of Bindra’s term ‘gnostic assembly’ or any other form of cognitive organisation necessary for object recognition, while ‘representation’ I will try to reserve for cases of particular mental images or memories of events which are assumed to direct sensory activities, searches for objects, or goal-seeking actions. The distinction is not always easy to uphold, and ‘inner description or representation’ abbreviated ‘IDR’ is a useful reminder that the shorter word ‘idea’ covers much the same ground.

The naming of hypothetical mental states or properties is of course a dangerous step, because, for people if not for animals, the use of a name gives rise to a sense of reality, and the ‘thinghood’ of what the name refers to, that may turn out to be spurious. Thus, little importance should be attached to the names themselves; what matters is that, in order to account for cognitive modes of perception, which are measurable in terms of reactions given to certain sets of external stimuli, it is necessary to posit the existence of forms of analysis of picked-up or received sensory information which might just as well be called ‘central cognitive structures’ or ‘schemata’ (as they are by Neisser and Piaget) or ‘small-scale models of external reality’ (as they are by Craik, 1943), as inner descriptions and representations. In other contexts, it is apt to refer to similar facets of mental organisation as ‘memories’ or ‘expectancies’, emphasising that mental representations of events can be tagged as past and future, to the extent that we might be able to deduce the existence of an apprehension that ‘something important is going to happen soon’ or a sense that ‘it’s been a long time since anything happened’ with only the vaguest knowledge of particulars.

However, the main point is that it is the capacity for complex perception, exemplified by visual pattern recognition of variable and disparate forms of the letter ‘A’, which requires us to consider these perceptual descriptions or schemata. Could it be the case that these complex forms of perception are a property of only the human mind, so that there is no need to assume that the necessary kinds of mental organisation exist in other animals? In previous sections I have argued that animal perception is cognitive, in so far as attention to such stimulus dimensions as pitch and colour, and to cross-modal qualities such as distance and spatial position, is free from particular behavioural reactions, and selective attention of this kind can be seen to vary. - But perhaps this is true only for relatively simple forms of perceptual analysis, so that the perception of certain colours, tastes, smells, sounds and rudimentary features like touched
corners or seen dots, curves and straight lines, may be subject to context and selective attention, without being capable of being formed into the more elaborate descriptions needed for human perception. Animal perception is often discussed as if this were the case: and certainly, just as there is much in animal (and human) reactions to sensation that is reflexive, there are many ways in which very simple cues of brightness, temperature and so on, influence animal behaviour. But consider how it comes about that a pigeon recognises another individual (as occurs for instance, in parent-young and parent-parent social activity). Or how a dog recognises particular human individuals. Although more direct methods, such as the detection of scents by dogs, may have their uses, it would considerably expand the adaptiveness of animals’ perception if objects could be recognised by assessment of the same range of sensory evidence that is surveyed in human pattern recognition. Also, the anatomy and physiology of the sensory organs and associated brain systems in mammals suggests they possess the same sort of physical machinery that accomplishes the schemata and inner representations of human perception. Perhaps we ought not to rely on these hints and suggestions, but fortunately it is possible to obtain additional evidence from behavioural tests.

Experiments in two-way classification and visual concepts

In fact we can look first at evidence that the pigeon visual system accomplishes the recognition of the letter ‘A’ by feature analysis of the same sort as was put forward as a theory for the human ability. The experimental technique which provided this evidence is not necessarily ideal, but is used very generally: it is one in which animals are trained to sort stimuli according to a two-way classification. For the animals, the classification is based on a ‘food’ versus ‘no-food’ dichotomy. Probably this makes the type of perception involved something of a special case, in that the sort of looking done in the context of food-seeking may differ from that done in, for instance, the homing or social interaction of birds. Manipulating the availability of food is, however, an invaluable method for the experimental study of animal capacities, and if unusual perceptual acts not apparently connected with the animal’s normal food-seeking behaviours can be thus observed, this is all to the good in demonstrating the flexibility of the perceptual mechanisms.

In the experiment reported by Morgan et al. (1976) pigeons were initially trained to look at 40 slides showing the letter ‘A’, and 40 slides showing the numeral ‘2’. These were back-projected one at a time, for an average of half-a-minute each, onto a small screen. Unusually, the experiment was conducted on free, rather than captive, birds, so the screen was in an apparatus presented at an outside window, rather than inside a small Skinner box. Several fantail pigeons living wild in Cambridge voluntarily took part in the experiment because, by pecking the illuminated screen, they could
produce food from an automatic dispenser. However, in accordance with standard laboratory procedures, food was only forthcoming if they pecked the screen while a letter ‘A’ was shown, and was not available when the numeral ‘2’ was shown. (In this instance, the birds were allowed to peck at grain for three seconds at the end of every presentation of an ‘A’, provided they pecked first at the screen.)

The results, for the three pigeons studied, were fairly straightforward: these birds pecked 2 or 3 times per second at the ‘A’s, and infrequently or not at all at the ‘2’s. In a sense this is unremarkable, since a two-way classification of ‘A’s and ‘2’s is not the same as the identification of ‘A’s from all other letters and similar shapes. There are a number of distinct clues, any one of which will differentiate between ‘A’s and ‘2’s, for instance a curve as opposed to a point at the top, or legs as opposed to a horizontal line at the bottom. However, I have not yet mentioned a crucial point: in the 40 slides of each category (‘A’ or ‘2’), eighteen different typefaces (from commercial designs) were used. Thus, although all the ‘A’s were capitals, some had rounded tops instead of pointed tops and there were variations in the size, thickness, and style of the lettering. Furthermore the introduction of 22 new typefaces, after the initial learning, did not cause any disturbance of the discrimination between ‘A’s and ‘2’s. Therefore any individual features which might have been used by the birds, such as the tops or bottoms of the figures, needed to work over a considerable range of precise forms of retinal stimulation.

To get a better idea of exactly what aspects of the lettering were being picked up, two further tests were made. Bits of ‘A’s and ‘2’S, and upside down and partially rotated whole ‘A’s and ‘2’S were shown to the first three birds in the same way as the normal stimuli, and two other pigeons who learned the original discrimination were then shown one at a time, a complete set of the letters of the alphabet (in the upright capitals of a single script). The results of these additional tests are revealing, although not absolutely clear. The birds were keen on an upside down ‘A’, and the top or bottom of an ‘A’ separately, but not on an ‘A’ on its side, or on the separate angled sides of an ‘A’. Of the other letters of the alphabet, R, H, X and K were responded to almost as if they were ‘A’s, and E, S, C, G, I, J, Land Z were treated as if they were ‘2’s, and not responded to at all. W, N, M, B, and U got some attention, in decreasing amounts, with progressively fewer responses being given to Y, T, F, D, V, O, P and Q.

What may we conclude from all this? It looks very much as though ‘feature detectors’ for ‘apex’, ‘legs’ and ‘horizontal line in middle but not bottom’, were utilised, with the presence of any one of these features eliciting some responding. Enclosed regions and T-junctions, which are also present in ‘A’s, did not seem to be picked up in isolation, since P and E were a long way down the list. ‘Curved top’ and ‘flat bottom’ which are present in ‘2’s, may have
been counter-indications to responding. Although this evidence by itself is rather limited, it points firmly in the direction of a form of perceptual analysis in the bird visual system more like a general purpose pattern-recognition device, and less like a restricted set of reactions to peripherally built-in cues such as brightness, colour and movement.

The reader may experience some reluctance to take seriously the notion that the pigeon eye and brain should recognise letters in anything like the same way as the human eye and brain, since, after all, what would be the point of a bird evolving an ability to deal with letters, and surely the human eye must have special ways of reacting to the printed word, since we all do so much reading. There is something wrong with this objection. The human visual system was not designed so that it could recognise letters; rather, letters were invented because they could be processed by the human visual system. Speech and language go far enough back to be specially built-in to the human brain, and that may be why we take the trouble to read and write, but reading and writing themselves are comparatively recent developments, and their visual components have to be accomplished with an eye and brain evolved for a life of illiteracy.

If, on the other hand, one is tempted to argue that discrimination between ‘A’s and ‘2’s is too simple an achievement on which to rest the attribution of perceptual schemata to pigeons, consider some of the other two-way classifications which these birds can apparently make when shown pictures by means of a slide projector. In the original experiment of this type (Herrnstein and Loveland, 1964), it is extremely difficult to suggest individual visual features such as ‘apex’ or ‘curved top’ which could have served to simplify the pigeon’s perceptual task, because the two categories of visual scene were colour slides showing people, versus colour slides not showing people. Over 1,000 slides of town and country landscapes were used, half of them containing one or more persons, in foreground or background, standing or sitting, clothed or unclothed, and often partly obscured, with the other half similar scenes containing no people. Each day birds were shown 80 slides, one at a time, for roughly a minute per slide. In this experiment, the conventional tactic was used of putting the birds inside a box, with slides rear-projected on a screen a couple of inches square on one of its walls. Of the 80 slides, half would contain people, and the other half not, these two categories always being mixed in a random order. The bird’s concern was to activate a food hopper, which it could do by pecking a hinged switch to one side of the projection screen. The bird did not, in this case, have to peck the screen itself: such factors often make a big difference to the observed sensitivity of animals to external stimuli, since their attention is sometimes narrowly focused in the direction of their actions, but experiments such as this one show that pigeons need have no great difficulty in looking at one place before pecking at another.
The critical aspect of the experimental arrangements, readily detected by the birds, was that as long as one or more persons were present in the view projected on the screen, there was a small chance that pecking the adjacent switch would trigger the food hopper, allowing them access to grain for a few seconds. The odds were such that on average the hopper could be operated just once for each slide with a person in it, but the exact time when this would happen was unpredictable (a ‘variable interval’ schedule of reinforcement: see pp. 92ff.). When no person was shown on the screen, food was never available.

As would be expected if information was more direct (by such categories as triangles for food and circles not), the birds, after a week or so of training, responded more vigorously when the slides which did contain people were displayed, and after a few more weeks made a very clear discrimination between the two categories. For almost all the slides with people, their reaction was measurable by the fact that they pressed the switch about 50 times during the minute each was presented. By contrast, for the scenes without people, the pigeons did nothing, or made fewer than 10 pecks.

Each day a random selection of slides was made from the collection of 1,000, so the birds had no opportunity to learn very specific stimulus features. There was thus no doubt that the 1,000 or so visual displays were being put into two categories, and the most obvious interpretation is that a sufficiently complex inner description, or visual concept, of ‘people’ was being utilised. Experiments using almost identical techniques have established that similar two-way classifications of slides can be made, by pigeons, according to whether the slides are of pigeons or other birds (Poole and Lander, 1971); of oak leaves or leaves from other species of tree (Cerella, 1979); of scenes with trees or similar scenes without trees (Herrnstein et al., 1976; Herrnstein, 1979); of scenes with versus scenes without bodies of water (Herrnstein et al., 1976); and of scenes with versus scenes without a particular human being present (Herrnstein et al., 1976).

In my view, the theoretical implication of these results is that the normal mode of visual perception in birds is the formation and comparison of inner descriptions or perceptual schemata. What I mean by this is that the birds use their eyes to perceive things, rather than using them as collections of isolated detectors for qualities such as brightness and colour. This is not saying a great deal, though it is useful to have experimental support for the ability of birds to do more than react to simple inbuilt visual signals. Many more detailed theoretical questions remain to be answered. Herrnstein et al. (1976) suggest that the visual capacities, hence inner descriptions, exhibited by the birds have ‘something to do with evolution’ and that the object categories for trees and bodies of water are ‘somehow represented in the genes’. But there was very little evidence to justify this, since the birds classified slides according to the presence of ‘trees’ or ‘water’ no more readily than they classified slides according to the presence of ‘people’, or according to the presence or absence of a particular young woman. It is not very
sensible to suggest that domestic pigeons come equipped with a complete set of perceptual schemata for all the things they are likely to encounter in the modern world. Whatever natural predilections or preferences may be characteristic of the species, it seems to be the case that the pigeon brain is capable of rapidly forming descriptions of a very wide range of visual patterns.

In the phrase used by Cerella (1979), ‘concept formation is spontaneous rather than deductive’. This refers to his result that a generally applicable description of some types of pattern can be formed on the basis of one example, and does not have to be distilled from a wide variety of experiences. At any rate, this happened with the outline of an oak leaf in Cerella’s experiments, in that the training given, of classifying one oak leaf against leaves of numerous other trees, or even, in one experiment, experience of pecking at one example of an oak leaf outline, without the benefit of other comparisons, enabled pigeons correctly to classify new oak leaf examples. Not too much should be made of this, perhaps, since one oak leaf outline looks very like another, by the standards of variability which apply to pictures of different individual human beings in different postures and from different angles, but it serves to emphasise that the original point of talking about ‘descriptions’ was that the descriptions should be sufficiently general to cover a variety of different instances, and, in the case of ‘A’s, several quite different forms.

The amount of experience needed to establish perceptual schemata of different visual categories ought therefore to vary to some degree according to the complexity of the category. There has as yet been little systematic investigation of this possibility. It is a matter of laboratory experience that simple discriminations, such as black versus white, for any seeing animal, and red versus green for animals such as the pigeon with strong colour vision, are formed more readily than discriminations which we should expect to be more difficult, such as a differentiation between two similar shapes of grey, or between two lines at slightly different orientations. It is common practice to distinguish between ‘easy’ and ‘hard’ discriminations, and to suppose that laboratory animals required to perform a hard discrimination will learn to do this more quickly if they have previously mastered a related, but easier form of it (see Mackintosh, 1974, p. 595 and Terrace, 1963). One would think, therefore, that it would be possible to demonstrate, because of prior experience or inherent complexity, that, for instance, pigeons can distinguish between individual pigeons more easily than between individual humans, or that they can distinguish between the two species more reliably than they can distinguish between individuals of either species.

Little evidence pertinent to this topic is available. Herrnstein (1979) suggests that, rather surprisingly, pigeons learn to notice the presence or absence of trees in visual displays just as quickly as they learn to notice the presence or absence of a loud tone, and just as quickly as they distinguish between closely similar colours. By comparison, although aerial photographs showing man-made
objects are successfully sorted from those not doing so (Lubow, 1974), pigeons are reported to be relatively poor at classifications of line drawings (Cerella, 1975) or at detecting the presence of object categories such as bottles or vehicles in photographs (Herrnstein, 1979). The precise nature of visual information which can and cannot be satisfactorily classified by pigeons therefore remains to be determined. Very little is known of how other species (even monkeys) might perform on similar classification tasks. A theoretical possibility is that the inner descriptions of visual information may not be confined to visual analysis. If, for instance, there is an evolutionary basis to the visual categories of trees and water, as Herrnstein et al. (1976) suggest, one would suppose that the description of ‘trees’ should include some estimate of value, and association with the possibilities for perching, and that ‘water’ should be even more definitely valued, and should have associations with taste and the actions of drinking.

**Abstract qualities in vision—regularity and irregularity of configuration**

The advantage of the two-way classification experiments discussed above are that the results are easily summarised and give clear support to the view that animal perception allows the detection of categories of reality which correspond to what we ourselves comprehend as natural or unnatural objects, rather than being limited to the transmission of isolated elements of sensation. The experiments may be less than conclusive, but the results fit well with theories such as that of Craik (1943) to the effect that perception in general must ultimately be understood in terms of its function in providing inner representations which parallel external events. It is unfortunate that data from species other than the pigeon are not available, but the study of primate cognition, which will be discussed in a later chapter, provides a wealth of evidence that inner descriptions of object categories are not confined to humans and pigeons.

There are, of course, many other forms of experiment by which the nature of animal perception can be assessed, but these are subject to great difficulties of interpretation. In between experiments on the detection of simple stimulus qualities such as colour and brightness, and the attempts to demonstrate the utilisation of visual concepts at the level of object recognition, there are some experiments which imply the operation of intermediate levels of perceptual analysis. At first sight, the traditional ‘oddity problem’ can be put under this heading. If a chimpanzee is several times shown two cups and a spoon, placed in a line, and rewarded for picking up the spoon, whatever position it appears in the line, it would be very foolish if, on first being shown two spoons and a cup, it did not pick up one of the spoons. However, if, over a longer period of experience, it is always rewarded for picking the cup from a line of two spoons and a cup, but the spoon from two cups and a spoon, it may very well adjust its selections according to this rule. If it did so, some would undoubtedly say that it had ‘solved the oddity
problem’, since in each case it would be picking what in fact was the ‘odd one out’ of the three objects. But the evidence as it stood then would allow us to infer very little about any inner conception of ‘oddity’ by the animal. It might have learned only to pick up an unduplicated spoon and an unduplicated cup, without being able to extrapolate from this experience when confronted with two saucers and a knife. This would involve, certainly, the perception of duplication in the case of cups and spoons, but without the extraction of a general principle that isolated objects should be picked up and duplicated ones left alone. Very careful experimental designs are required to decide how specific particular perceptual strategies are. However, with visual stimuli, and one odd out of three stimuli presented at once, it is probably the case that a much more general rule of ‘oddity’ is perceived by chimpanzees than by pigeons. That is, chimpanzees who have been trained to select a red disc from two blues and a red, but the blue one from two reds and a blue, continue to apply the rule of oddity if the objects presented are changed to triangles and squares. (Strong and Hedges, 1966; Bernstein, 1961). Pigeons require prolonged training even to perform correctly in picking out the odd colour, when two colours are unevenly distributed in three patches, and show few signs of having grasped a principle when tested with a new pair of colours, let alone after a shift from colour cues to shape cues (Zentall et al., 1974; Urcuioli, 1977).

Rather surprisingly, it is very difficult to interpret such apparently obvious differences in perceptual capabilities between species. Perhaps the first hypothesis would be that mammals, with brains much more like our own than birds, should perceive visual stimuli in terms of more abstract rules than birds. But mammals other than chimpanzees and monkeys, even cats and racoons, have shown little sign of a clear superiority over pigeons on oddity problems (Strong and Hedges, 1966), and some species of birds other than pigeons may demonstrate monkey-like levels of abstraction on visual tasks— canaries (Pastore, 1954) and rooks (Wilson, 1978) seem to generalise oddity from one set of visual cues to another.

Although a difference between chimpanzees and pigeons is not necessarily an instance of a universal gulf between birds and mammals, it may represent the superiority of primates over all other animals, and this possibility will be considered in a later chapter. For the present I want to examine not the way in which modes of visual perception may vary from species to species, but aspects of the analysis of the visual field which may to a large degree be shared by a great variety of vertebrate species.

Some characteristics of the oddity problem point to a rather general property of vision, and that is the detection of irregularities and discontinuities in patterns of retinal stimulation. Even within the retina, and at the lowest levels of any hierarchy of visual analysis, it is contrasts and differences between adjacent parts of the visual field which serve as the units of stimulation. At some subsequent level, it
seems likely that discontinuities in the visual field as a whole attract special attention. A single dark spot set against an otherwise blank background has much more subjective significance, to the human observer, than a dark spot in the same position which is only one among a multitude of similar spots scattered over the visual field. Is this a product of human awareness; or is human subjective experience in this case a product of forms of visual analysis which are shared by other possessors of the vertebrate eye? And what if our attention is drawn, not to a single spot against a featureless background, but to a single large spot in a scattered array of small spots, or a single cross in a regular array of noughts of the same size? Should we assume that a pigeon, or a cat, is subject to the same kind of perceptual effect?

Any attempt to explain how the visual system works in species other than ourselves would have to consider this possibility, and there is a certain amount of evidence from behavioural experiments to suggest that the mechanisms which underlie many subjective effects in human visual experience, such as the ‘figure-ground’ effect by which one item stands out from a background, have counterparts in other species.

For instance, the oddity problem itself can be conceived of as a test of the ‘figure-ground’ effect—how easy is it for animals to pick out the odd item out of a collection of repeated elements? This question applies to any particular set of elements—clearly only species with good colour vision will-pick out a figure of one colour from a ground which is of a different colour, but equally bright. And a species which is very sensitive to discontinuities in colour may not be able to respond to triangles hidden in an array of squares. If we were sure that both these perceptual effects occurred in the same species, it would still be an open question whether an animal of this species trained to direct a response at a discontinuity of one type would feel obliged to make the same response when first presented with a discontinuity of the second type.

A confirmation of the figure-ground sort of perceptual effect within a conventional oddity detection task performed by pigeons has recently been obtained by Zentall et al. (1980). Pigeons were shown a row of three adjacent discs, each about half an inch in diameter. When two discs were red and one green, pecking at the green one was rewarded; when two were green and one red, pecking at the red one was rewarded. The birds failed to learn this task when the two conditions were randomly alternated. There would of course be no problem for the birds in always picking out the same colour, irrespective of its position in the row—it is having to ignore any preference or learned rule of pecking at a particular colour which makes the problem difficult, and allows in this case the assessment of the extent of the contrast between duplicated and non-duplicated colours.

The most revealing test in the experiments of Zentall et al. (1980) involved presenting pigeons not with a row of three discs, but with an array of 25 of them (in a five by five matrix) with the
same alteration as before: either one disc was red with the other 24 green, or vice versa. In this case the animals quickly learned to peck only at the ‘odd’ stimulus wherever it appeared, with almost 100 per cent accuracy. The obvious implication is that the perceptual contrast between one odd element and an alternative which is very extensively duplicated is much more detectable than the contrast between a single element and an alternative only duplicated once.

Especially with colours, it is easy to imagine this as an immediate and automatic effect, similar to the perception of a black dot against a continuous white background, or a white spot against a black background, and an aspect of automatic methods of processing of visual information. That colour is not the only type of visual information processed in this way is indicated by some results used by Sutherland to support his theory of structural descriptions in animal vision.

First we may consider an experiment on the perception of regularity in black and white checkerboard patterns by rats. Although some rodents, such as squirrels, have good vision, the rat as a species is not noted for the acuity of its eyesight. (Furthermore albino varieties of laboratory rats, like other albino mammals, have, as well as difficulties due to lack of pigmentation in the eyes, oddities in the visual pathways in the brain.) Visual discriminations can be obtained in rats, however, especially if a rule of ‘look before you leap’ is engaged by forcing the animals to jump towards placards on which various shapes and patterns are drawn. By offering the animals a choice of two placards to jump to, a kind of two-way classification of visual displays can be obtained. In the experiment reported by Sutherland and Williams (1969) rats (not albinos) were given the choice of either jumping towards a completely regular ‘checkerboard’ pattern (of 16 black and white squares), or jumping to a similar display which contained a ‘mistake’ in the form of a displaced pair of component squares, so that there was a row of 3 black squares over 3 white ones at one place in the array.

If the rats were rewarded only for jumping to the regular pattern, then, characteristically, they took a long time to manifest accurate visual discrimination, taking about 2 weeks, at 10 choices per day, to reach a criterion of not more than one wrong choice per day. Now one’s initial hunch about the visual processes taking place during this period would be that the animals were learning to look for the ‘row’ of black or white squares in a particular place. Something of this kind probably did occur, since irregularities in the lower half of the display screens, where rats tend to concentrate their attention, were picked up soonest. However, this was not the only thing going on, because, after they had other experience of three sessions of an irregular pattern, the rats managed to sustain their correct choices when new instances of an irregularity were introduced.

Sutherland’s conclusion from this was that rats form and store highly abstract rules describing the pattern of visual input which they are exposed to. One could argue about exactly what ‘abstract’ means in this context, but it is clear that something about the
patterns, in terms of regularity and irregularity, is abstracted from the visual arrays, something which is independent of the exact sequence of light and dark squares. Quite possibly, the analysis of light and dark squares is related to the analysis of the array of red and green circles by the pigeons in the oddity experiment of Zentall et al. (1980), discussed above. Subjectively, certainly, there is to a human observer something immediate and automatic about the way in which an irregularity in a checkerboard pattern stands out from the rest, just as one picture hanging askew on a wall demands attention. Conceivably, part of the aesthetic appeal of repeated motives in wallpaper patterns, friezes and similar forms of decoration, is that the regularity and symmetry is perceived as a quality of visual input. It is unlikely that laboratory animals care very much about regularity and the symmetry of repetition in this sense, but the Sutherland and Williams experiment suggests that rats are capable of noticing regularity, at least as a property of checkerboards, and Delius and Habers (1978) have shown that pigeons can classify shapes into categories of symmetrical and non-symmetrical about the vertical.

A second experiment demonstrates a similar distinction between regular and irregular figures, but is notable because the subjects were goldfish, which, although they have quite mobile and sensitive eyes with better quality retinas than the rats’, might not be expected to have enough brain circuitry beyond their eyes to do more than detect exact repetitions of particular patterns of retinal stimulation. But, according to the data collected by Bowman and Sutherland (1970), this is far from being the case. Visual perception in the goldfish may be studied successfully by a number of methods, but one of the most direct is to obtain two-way classifications by eliciting the choice of one of a pair of figures shown to the fish at the same time. In this experiment the fish were shown a perfectly square piece of black plastic sheet, and a square identical except for a bump on the top, at the end of their tank, one shape being baited with food, and the other not. Under these conditions, goldfish very rapidly demonstrated their visual acuity by learning to swim to the shape at which they had previously found food. (Some were trained with the ordinary square baited, others with the bumpy square baited.)

The question is, what mechanisms of analysis of visual information allow for such a discrimination? The bumpy square presented a rather greater area— is it discriminated on the basis of greater total area, or possibly by the presence of an additional area of blackness in a position at the top of an ordinary square? Such hypotheses can be tested by giving trained fish new pairs of stimuli which they have not seen before. If only total area were detected, then a smaller bumpy square or a square with a chunk carved out, might be confused with the ordinary square. In fact no such confusions were observed. The rule apparently being followed by the fish was that any square with a sudden break on the side where the bump was on the original bumpy square, should be treated as a
bumpy square. In a rather subtle extra test, the fish were shown the ordinary square, but with an extra vertical strip of plastic just in front of it, giving a two-dimensional outline like the bumpy square. If only the outline of the pattern of illumination impinging on the retina was important, the goldfish might have been forgiven for responding to the compound stimulus as if it was bumpy. But this did not happen: a square with the site of the original discontinuity occluded by the overlaid strip was responded to randomly, as if (quite rightly) it might or might not have had an irregularity at the hidden place.

The whole experiment (Bowman and Sutherland, 1970) was more elaborate than this, with a number of forms of the original irregularity, and of subsequent test figures, presented to different fish. All the results, however, were consistent with the conclusion that the information picked up from the original displays was ‘highly abstract’, at least by comparison with the supposition that the only capacity contained within the eye and brain of the goldfish is the connection of exact patterns of retinal stimulation with particular swimming movements.

That goldfish trained to swim to a square with a small triangular extension on the top side will also swim to a circle with a small semicircular bite out of the top, might be thought to indicate only a certain vagueness about their vision. It should be emphasised that their acuity is not in doubt, and that they are perfectly capable of picking out circles from squares, and bumps from gaps, if the occasion requires it. Presumably there are limits on the complexity of visual descriptions that can be utilised by fish, and we should expect these limits to be set very much lower than those for the pigeon, and for the nature of their perceptual schemata to be quite different from those of the chimpanzee. Behavioural evidence for such species differences is difficult to obtain, but it is a step of some importance to discover that even a vertebrate as small and as psychologically insignificant as a goldfish appears to subject visual information to such varied levels of analysis.

**Comparative anatomy and physiology of the visual system**

The behavioural evidence suggests that goldfish can detect a visual category or quality of ‘irregularity’ over a wide range of immediate physical stimuli, and that pigeons can combine features which may have a similar degree of remoteness from particular patterns of light into classifications which approach the complexity of a description which will allow the visual identification of a letter of the alphabet, or of human presence. Unless such evidence is seriously flawed, it implies the availability of biological machinery appropriate for these tasks.

In fact, only if one held simple-minded views about evolutionary scales of excellence would one doubt whether lower animals like the
goldfish and pigeon had *eyes* which were sufficiently developed to serve in the recognition of complex visual patterns: we may still be surprised that their brains are capable of dealing with the richness of the information supplied by their eyes, but this may be a case where experimental studies of the psychology of animal perception forces a re-evaluation of the theories of brain function.

Although there are enormous gaps, there is sufficient knowledge of the basic anatomy of vertebrate vision to allow for some comparisons to be made between the machinery itself, and the way it works (see, Walls, 1942; Rodieck, 1973; Kruger and Stein, 1973; Masterton and Glendenning, 1979). The first point is that the eye itself displays more uniformity in structure, from species to species and from class to class, than any other organ of the body. In terms of the peripheral origin of sensation, therefore, human vision is less distinguishable from that of other vertebrates than is the case with other sensory modalities. For touch, for instance, the difference between skin and scales gives no common starting point to the perceptions of fish and mammals. And while the hair cells which convert vibrations into nerve impulses are essentially similar in all vertebrates, the development of sound-transducing organs is quite different in fish and air-living classes, and there are radical changes in the structure of the ear between amphibians, reptiles, birds, and mammals.

The eye of all vertebrates works like a camera, in that light reflected from objects is passed through a transparent lens to be focused on a sensitive film of tissue a short distance behind. In fish and amphibians the lens is moved backwards and forwards for focusing, as in a camera, but in reptiles, birds and mammals, the lens is thickened or made thinner by the tightening or relaxing of muscles. The very first visual distinctions are thus those between crisp and fuzzy images, fed back into focusing mechanisms. The light-sensitive tissue at the back of the eye, the retina, works in a more or less similar way in all vertebrates, although there are certainly plenty of differences from species to species. Receptor cells which are most sensitive to dim light (‘rods’) are distinguished from those somewhat less sensitive, but useful for high acuity, and for differential response according to wavelength, as the first step in colour vision (‘cones’). Several other types of cell, within the retina, intervene between the detection of light by the receptor cells, and the ‘ganglion’ cells which provide the output from the retina to the optic nerve, and which are thus the point at which information from the eye is passed on to the brain.

There are two points of some controversy here: first, how complicated is the information that retinal ganglion cells send back to the brain in different species, and second, how much does the brain act back on the ganglion cells or other parts of the retina, to tune them selectively? But an agreed general concept is the ‘receptive field’ of cells (Kuffler, 1953), meaning the patch of the retina which, when stimulated by an appropriate pattern, will make a cell fire off nerve impulses. The typical type of pattern which
activates retinal ganglion cells is a small spot of light surrounded by
darkness, or black dot surrounded by light (‘centre-surround fields’). These have been measured most clearly in the cat and monkey, but are also apparent in goldfish. In frogs, pigeons and rabbits, somewhat more complex visual events seem to be detected within the retina, such as lines or dots moving in a particular direction. This has given rise to the idea that mammals such as the cat, the monkey and man have rather simple retinas, waiting to analyse shape and movement until the ‘on/off’ sort of information has reached the brain, whereas less complicated animals get a greater amount of feature analysis over with before transmitting messages down the optic nerve. De Monasterio and Gouras (1975), however, who found some movement sensitive cells in rhesus monkey retinas, suggest that this distinction may have been overplayed. In the human retina, about 125 million primary receptors (rods and cones) converge, through the intervening network within the retina, on less than 1 million ganglion cells, so a good deal of summarising must go on in this case. Subjective detection of moving lights occurs for such rapid movements that Wertheimer and McKee (1977) think movement detectors in the human retina are a possibility, and the separation of different types of information according to different sorts of fibres in the optic nerve, in monkeys and cats, has become more and more apparent (Lennie, 1980).

For most intents and purposes, the retina can be considered to act independently of the brain. There is some evidence, however (e.g. F. A. Miles, 1970), that in pigeons and chickens the brain may act forward to the retina to maintain the sensitivity of certain ganglion cells. It is sometimes suggested that similar modulating of signals from the retina, by outgoing impulses from the brain, occurs in other species such as the cat and monkey, but whether this is a very important influence on the way the retina transforms the optic image is doubtful (Rodiek, 1973).

**The relative importance of vision in vertebrate evolution**

It is fairly safe to say that the eye, and in particular the retina of the eye, is a very sophisticated device for transforming optical images into activity in the optic nerve, in all classes of vertebrate, and in the goldfish, pigeon, rat and monkey in particular. Many (though not all) of the psychologically interesting questions have to do with what happens to neural activity at the other end of the optic nerve. A peculiarity of visual perception is that the usual phylogenetic hierarchy of brain function, in which we should expect that the brain of the monkey has more in common with that of the rat than with that of the pigeon and goldfish, conflicts with the similarities of the visual input to the brain in fish, bird and primate. The similarities are (a) colour vision, and (b) foveal vision: they may be deceptive, and are probably best explained in terms of their absence in mammals other than primates.

The ‘bottleneck’ theory of the evolution of vision in mammals (Masterton and Glendenning, 1979) has it that early mammals were nocturnal, and thus less sensitive to colours. Living in complete
darkness may mean that species lose vision altogether (as in the case of moles and blind cave-living fish) relying on touch and/or smell. Occasionally species evolve spectacular alternatives to vision, as in the echo-location systems of bats and dolphins and the electric field method of some species of fish living in very muddy rivers. It is therefore not implausible that early mammals sniffing around in the dark for roots and grubs might have de-emphasised vision in general. To the extent that vision was retained, in these nocturnal species, sensitivity to dim light, rather than acuity, would be at a premium and therefore a retina composed largely of rods would have been necessary, and hence colour vision would be lost. Finding safety in trees during the day, instead of prowling about at night, and specialising in a diet of fruit, may be supposed to have led to the invention of acute colour vision, with cones as well as rods in the retina, in only the primates among mammals.

Apart from sensitivity to colour, one of the crucial aspects of the human eye is the use of a fovea, and an ‘area’ of the retina. The ‘area’ is a circular patch in the middle of the retina which has a high concentration of cones, giving high acuity, and the fovea is a depression in the middle of this, which leaves only a thin covering above a small spot with no rods and a high density of closely packed cone receptors. The fovea is what we look at things with, and the use of this place on the retina is crucial for detailed human vision. Characteristically, human vision is to a large extent foveal vision. (It has been claimed, for instance, that reading is only possible when the images of words fall on the fovea: Rayner and Bertera, 1979.) No mammals apart from primates have foveas (though some, such as the cat, have central areas where rods or cones are concentrated). But many species of marine teleost fish, and of reptiles and birds, do have foveas in their retinas. It is not uncommon for birds to have two fovea in each eye, and on structural grounds, birds can be said to have ‘better’ fovea—deeper and steeper pits—than man or the monkey. The duplication of foveas in the eyes of bird arises because of the relative immobility, and lateral placements, of the eyes. Foveas in the centre of the retina face to the side of the bird, and are therefore used for accurate fixation at each side, but foveas at the outer sides of the retina allow for binocular inspection of objects in front of the bird. Hawks, eagles, swallows and terns are among the birds with two well-developed foveas in each eye; more common is one central fovea, as in the pigeon or sparrow. As well as the central fovea, the pigeon has a special ‘pecking field’ in the upper side quadrant of each retina, which is distinguished by red oil droplets as filters for the cones. Some birds (such as gulls) and some mammals (such as rabbits), have a horizontal stripe of densely packed receptors, thought to be used for fixation of the horizon.

To the extent that the phenomena of human vision depend on the fovea itself, therefore, some human specialisations might be shared with birds: more generally, vision in particular species must depend on the characteristics of the retina and comparisons between species must take this into account. It can usually be taken for granted that a predominance of cones over rods in the retina is backed up by brain
mechanisms which compare the output of different types of cones to produce sensitivity to colour. The legendary acuity of birds of prey can be understood in terms of the excellence of their retinas—the fovea of a large hawk (*Buteo buteo*) is packed with cones at a density eight times greater than that found in the human fovea, and, as in the human case, almost every receptor is represented in the optic nerve. Even the non-foveal parts of the retina of these birds would be expected to have twice the resolving power of human acute vision.

Thus the information which the eye sends to the brain in non-

human primates is very similar to that sent from the human eye, and the amount of information derived from the eye in other species, especially birds, may match or exceed that which the human brain receives. For ourselves, we know that visual information is eventually transformed into the images and illusions of our subjective perception. How are the reports dispatched from the retina interpreted in the brains of other animals?

*The receipt of visual information in the vertebrate brain*

The complexity of this topic cannot be underestimated. The points I wish to make here are comparatively few. The main one is that our knowledge of the brain mechanism involved in human vision is based on evidence from animals, and in particular from cats and monkeys. In terms of physiological mechanisms, it is not at present possible to claim that human visual perception is in any way significantly different from that of the rhesus monkey or chimpanzee. However, there are apparently very radical differences between the brain mechanisms used for visual perception by mammals and those used by other vertebrates, differences already referred to in Chapter 5, and we need to return to the question of whether mammalian vision is therefore more cognitive than that of the other vertebrate classes, in the light of the behavioural data which suggests that goldfish and pigeons can form abstract and complicated perceptual schemata from visual experiences.

*Visual cortex; area 17; striate cortex—the mammalian super-retina*

The three terms ‘visual cortex’, ‘striate cortex’, ‘area 17’, are used interchangeably to refer to the critical arrival point of retinal messages in the mammalian brain (‘V1’ is yet another synonym). Perhaps as many as three-quarters of the fibres in the optic nerve relay in the thalamus (in the lateral geniculate nucleus; see Chapter 5 for thalamic functions) which sends the message of each of these fibres directly to the striate cortex, on the surface of the brain, right at the back of the head in man, and in a similar position in other mammals.

Great advances have been made in the last twenty years in the
study of the structure and function of this area of primary visual
cortex, and much has been made of what has been discovered
(Hubel and Wiesel,

1962, 1974, 1977; Berkley, 1979; Blakemore, 1975; Frisby, 1979). The basic technique of these investigations is to take electrical
crecordings of the activity of individual cells in the striate cortex of
lightly anaesthetised animals, while spots of light of different shapes
and sizes are moved over various parts of the retina (see Chapter 5).

Perhaps the most striking and fundamental fact about the
organisation of visual cortex is that it provides a point-to-point map
of the visual field. If we were to look at a large letter ‘A’ in front of
us, the pattern of activity in visual cortex would also form a letter
‘A’, though of a peculiar sort, since it would be upside-down, and
laterally reversed after being split middle to sides. In other words, if
we looked at a large poster saying ‘VISION’, someone looking at
the back of our head, if he could detect cortical activity, would see a
rough outline of the letters, all upside-down, in the order I, 0, N; V,
I, S. The semi-colon indicates that the letters I, 0, N, in the right half
of the visual field would be written across the left hemisphere while
V, I, S would be on the cortex of the right hemisphere. The
connections between the hemispheres for the striate cortex are
concentrated in the parts of the cortex that represent a straight line
right down the middle of the visual field (Zeki, 1978a). It is
legitimate, then, to think of the visual cortex as providing a ‘picture
in the head’, since topological relationships in outside two-
dimensional space are repeated on this surface of the brain. Apart
from being inverted and split, the cortical picture is stretched out in
the middle and compressed at the fringes, since space on the cortex
is allocated roughly according to receptors in the retina, which are
themselves more concentrated in the middle.

The current theory of what is actually going on in striate cortex
is rather involved (Hubel and Wiesel, 1977; Frisby, 1979) but can
be given a rough paraphrase. The striate cortex, on which the
transformed retinal image is displayed, is made up of a mosaic of
small blocks, each one or two millimetres square. Within each block
are columns of cells which respond to edges, slits (light against
dark) or lines (dark against light), of one particular angle or
orientation. These columns are arranged into the blocks (often called
‘hypercolumns’) so that a complete set of possible orientations is
covered. Within one hyper-column, there are different slabs for each
eye (in cats and monkeys, with front-facing eyes, everything is seen
with both eyes). Individual cells in any block can be found which
detect stationary edges, slits or lines in very precisely defined
positions in the visual field (‘simple cells’); other cells respond to the
movement, within small areas, in a

given direction, of an edge or line at a certain angle (‘complex’
cells) and yet others select lines moving over their area of the retina
which are of a certain preferred length (hypercomplex’ cells).
Thus it is as if the visual field is plotted out in squares, and a separate block of cortex assigned to each square of the visual field, so that for each square, the same set of questions can be asked—is there a light/dark edge in this square?—what angle is it at?—is it moving? and so on.

This is all good to know, and it provides confirmation of the physical existence of the lower levels of feature analysis discussed earlier in the chapter. But, as was emphasised then, these levels are only the most rudimentary beginnings of visual perception. The ‘hypercolumn’ theory tells us that if a particular form of the letter ‘A’ is flashed on a particular part of the retina, we can expect a particular set of firings of cells in the visual cortex. But it tells us almost nothing about how ‘A’s in different typefaces, at different positions in the visual field, are all recognised as ‘A’s. If the goldfish or pigeon retina were able selectively to respond to edges and movements in certain places in the same way (Maturana and Frenc, 1963) we would still be left with the puzzle of why the goldfish may swim to either a square bump, or a circular notch, as instances or irregularity, since different cells should be firing in these cases, to say nothing of the differences in firing patterns from the pigeon retina to slides of a clothed standing, or unclothed reclining person, both of which may elicit the same behavioural response (see above).

Since the striate cortex is a stage in the visual perception of mammals, and we would presume that mammalian vision is even further removed from reactions to individual edge movements than that of goldfish and pigeons, the fact that small regions of this cortex contain cells capable of detecting a wide range of visual features in corresponding small areas of the visual field does not take us very far towards understanding how inner descriptions, which could enable us to perceive objects, rather than edges, might be constructed. The features detected in the one thousand or so little blocks of striate cortex described by Hubel and Wiesel (1977) must obviously be fed into further stages of comparison, abstraction and interpretation. One question is simply how far these further stages are incorporated into the striate cortex itself, and how far they depend on other parts of the brain—in particular, areas of cortex surrounding the striate visual projection, conveniently termed the ‘extra-striate’ areas.
Facilities thus exist for an enormous amount of lateral interaction between closely adjoining areas of striate cortex (Fiskin et al., 1973). There is considerable sense therefore in the theory that several stages of abstraction and perceptual construction take place immediately, within the striate cortex itself. Even principles derived from psychological examination of subjective visual experience can be applied here. I have mentioned the subjective immediacy of the ‘figure-ground’ distinction. Similar configurational qualities described by introspective ‘Gestalt’ theories of the first half of the century were ‘proximity’, ‘similarity’ and ‘grouping’. At one Level, seeing continuous lines and figures and continuous movements may result from lateral filling in and synchronisation between adjacent sections of striate cortex. Seeing dots in groups of the same size, and noticing irregularities of colour and shape, could also be manifestations of lateral comparisons across the visual field as it is represented by cellular activity in this part of the brain (see Frisby, 1979, pp. 110—12, and Marr, 1976).

But even if the output from this primary visual projection in the mammalian brain takes a great deal of account of configurational aspects of patterns of light impinging on the eyes, it is unlikely that this output alone provides a useful description or schema for object recognition. Interactions within the striate cortex are almost certainly responsible for the difference between what we see with one eye alone, and what we see with both eyes open at once, but subjectively, under normal conditions, this does not amount to very much. Conceivably, a cup of tea two feet away, and the same cup ten feet away, although activating a much larger area of striate cortex in the first case than in the second, might produce patterns of striate activity with something in common, but subsequent stages in the visual pathway would be better placed to detect the commonality. For one thing, if the further away cup was in the right half of the visual field, but the closer cup was to the left, the initial descriptions would be in different hemispheres.

In any event, there is sufficient anatomical evidence to make it clear that many important aspects of visual perception occur outside the primary projection from the retina to the striate cortex. This is admirably summarised by M. Wilson (1979). In part the evidence can be understood in terms of what happens to visual information after the striate cortex has done its job of constructing a ‘primal sketch’ (Marr, 1976) from the optical image in the eyes. But directions as to what is to be looked for in the optical image may be given before the construction of any sketch. What is known about anatomical pathways suggests that there are cycles of perceptual analysis in which questions and answers are continuously passed round between numerous stages in the visual pathways, the striate cortex standing out because it contains the most faithful reflection of reports from the retina.
Apart from the horizontal passing of information across short distances within the Layers of cerebral cortex, the fact that the cortex itself is in a thin sheet means that there is plenty of space for outputs into the great tracts of white matter underneath the cortical grey, these tracts being made up of nerve fibres carrying impulses across to different surface regions, and down to the thalamus and other internal brain nuclei. The general plan of output from the striate cortex is that information is passed forward, from its location at the very back of the brain, first to the surrounding cortex of the occipital lobes, then further forward to the cortex in the parietal lobes, getting closer to tactile sensory projections, and further forward to the temporal lobes, getting closer to auditory projection areas. An advantage of having the middle of the external visual field located around the outside of the striate mapping is that the detailed and concentrated information from the fovea is right at the point where the transition to further stages of analysis takes place (Cowey, 1979).

Topographical maps of the visual field appear to be repeated several times after the initial striate projection, and it may be that each of these additional maps is specialised in certain features, such as colour comparisons, movement comparisons, left-eye/right-eye comparisons for depth perception, and so on (M. Wilson 1979; Zeki, 1978a, 1978b; Cowey, 1979). Because of the techniques of physiological investigation, cells which respond to certain classes of stimuli, in certain positions in the visual field, are the easiest to identify. But of even more interest would be regions in which position in the visual field is irrelevant, and cells do not respond consistently to particular simple visual features, because their job is to pick up more complex features, such as a discontinuity in an otherwise continuous line, or the occlusion of one object by another, over large sections of the optical image. The crucial thing about such higher-order analysers is that they must be optional—irregularities might be important in one context but not in another, and the main theme of the first part of this chapter was that analysers should be capable of being ‘switched-in’ or ‘switched-out’, according to circumstances. The technique of recording from individual cells in restrained and anaesthetised animals may not be as helpful for identifying these more flexible stages of visual analysis, which are important precisely because they are not inevitable responses to certain exact patterns of retinal stimulation. It is therefore hardly surprising that the use of these techniques has not allowed such detailed and systematic mapping of the functional arrangements of extra-striate visual cortex as has been possible for the more reflexive reactions of cells in the striate, ‘first-stage’ visual projection. Areas outside primary sensory projections have indeed often been defined as ‘silent’, since they usually do not manifest electrical activity in response to local stimulation at peripheral sense organs. However, when cells can be found outside the’ primary and secondary sensory cortex which respond to any sort of sensory input, the results are of exceptional interest, even if, and perhaps
especially if, the spatial arrangements of the cells do not fall into any obvious architectural plan.

The results obtained by Gross and his colleagues from the cortex in the bottom part of the temporal lobe in the brains of monkeys come under this heading (Gross et al., 1972; Gross, 1973; Gross et al., 1974). Cells in this area often respond to complex visual patterns, which may appear in wide areas of the visual field, and some cells respond if the patterns are presented to either the left or to the right visual field. Stimuli which are moved towards or away from the animal, so that the exact size of the image on the retina changes, are often very effective. Particular stimulus shapes such as a circle, a circle with a toothed

circumference, or a semi-circle, may be ‘recognised’ by particular cells, from retinal images of different sizes and at different positions. One cell was found which gave little or no response to any of these shapes, but fired vigorously whenever an outline approximating the shape of a monkey’s paw was presented. This region of the brain (‘inferotemporal cortex’), which is in fact closer to the main auditory receiving areas than it is to the striate cortex, may thus be the location of the more abstract parts of perceptual description, in the sense that it receives combinations of inputs which characterise objects, rather than specific and localised visual experiences (Gross and Mishkin, 1977). If the striate cortex exemplifies a ‘world of features’, it is left to other regions to perceive the ‘world of things’ (Wilson, 1979). It is the primate temporal lobe, which receives visual information rather indirectly, that is the ‘organ of categorization par excellence’ for seen events (Weiskrantz, 1974, p. 202).

Old wine in new bottles—the primitive tectum and the primate pulvinar

One reason for interest in the extra-striate cortex, which is beyond the main pathway from the mammalian eye to the mammalian brain, is that there is a relatively large -amount of it in human and primate brains. The part of the thalamus which relays to and from the extra-striate cortex is also enlarged in primates, and given a special name, the ‘pulvinar’. But, as with the other primate specialities of colour vision and the central fovea, this brings us back again to the sub mammalian classes including the fish and the birds. For the pulvinar receives some of its input, even in the highest primates, from the quarter or more of the optic nerve which splits off from the main mammalian projection, and goes first to the tectal areas of the midbrain (in mammals the main part is called the ‘superior colliculus’, but it is convenient to keep to the name of tectum).

Thus in primates some visual information goes from the eye to the tectum, from the tectum to the thalamus and from the thalamus to further projections in the cerebral hemisphere. Exactly the same thing is true of birds, except that a larger proportion of optic nerve fibres begin this route. Again the same thing is true of fish, although
the final forebrain destinations are better known in the large brains of sharks than in the smaller brains of the bony fish (Ebbesson, 1970; Réperant and Lemire, 1976).

Is this anatomically similar sequence serving the same function in primates as it is in the pigeon? Or, alternatively, has the sequence from tectum to forebrain been freed from its earlier role by the arrival of the striate region of mammalian cortex, remaining to acquire newer, more cognitive, functions? These two alternatives are clear enough, but it is surprisingly difficult to say which is true.

The first problem is that pigeons (to a greater extent other birds such as the owl or crow, and to a lesser extent all other non-mammalian vertebrates), while they may not have the striate cortex of a monkey or cat, do have a visual pathway from retina direct to the thalamus and then to the higher centres, which looks analogous to the main mammalian geniculostriate system (Nauta and Karten, 1970). Why does the pigeon need this if the larger eye-to-tectum is already doing what the striate cortex in mammals is supposed to do instead? The second problem is that the eye-to-tectum pathway, retained in mammals, seems to have some of the properties which it has in non-mammalian vertebrates (such as sensitivity to movement) and thus there is little direct evidence of its function being radically changed.

What can be said is that, in non-mammals such as the goldfish and pigeon, it is the optic tectum which is the most obvious ‘super-retina’. Each point on the retina maps on to a point on the optic tectum, which is a large and highly differentiated part of submammalian brains, with several laminated layers. In mammals such as the cat, with front-facing eyes, the projection on to the tectum is somewhat complicated, but it has a point-to-point correspondence with the visual field. However, in the cat, the new point-to-point correspondence on the striate cortex takes up much more of the available visual information.

The most convenient way to ascribe functions to these two internal mappings of the external visual scene inside the brain is to suggest that the older, tectal projection is for ‘noticing’ and the new and improved striate cortex projection in mammals is for ‘examining’ (Weiskrantz, 1972). ‘Noticing’ might require either vague knowledge of movement and brightness changes in a particular part of the visual field to be connected to various eye, head and body movements to bring the noticed source into better view, or to instigate reflexive attack or escape movements. ‘Examining’ on the other hand, might involve a much more detailed reconstitution of the optical image, with greater utilisation of previous experience to identify particular objects and object categories. Noticing would be more automatic and reflexive, with examining and identifying being more cognitive and controlled by context, as befits something especially well developed in mammals. The integrating of noticing and examining by the combination of striate and extra-striate mechanisms in primates
could be thought to be reflected in, for example, the cognitive control of eye movements, and the necessity to use a great deal of fairly reflexive noticing in the periphery of the visual field in order to direct fixation of the eyes for the foveal processes not present in other mammals.

To some extent, this distinction fits well with behavioural evidence, if we are content with the assumption that non-mammals do rather a lot of reflexive noticing, and relatively little examining. The apparently abstract coding of form irregularities by goldfish, for instance, could be a consequence of the vagueness of perception resulting from their retina-to-tectum projection. Possibly even the comparative generality of the two-way classifications of coloured slides by pigeons could be put down to rather global noticing capacities in their well developed midbrain visual analysis. There is a difficulty in that the behavioural evidence as it stands does not suggest that pigeons have any great difficulty with tasks that appear to involve some examining, such as finding the ‘X’ s in an array of ‘noughts’ (Blough, 1977, 1979) or selecting only the seeds which they prefer from a handful of mixed grain (in some cases eating wheat but leaving tares). The finding that they can classify together slides showing a particular human individual (Herrnstein et al., 1976) as well as recognising their own mate and young in natural conditions, also argues against any general lack of detailed form vision in pigeons.

Birds especially, among the non-mammalian classes, have a reasonable analogy to the mammalian striate cortex system, as well as massive projections from their midbrain tectal mapping to the thalamus, with further back-and-forth exchange of information between thalamus and higher centres roughly comparable anatomically to the extra-striate cortex circuits in primates (see Karten, 1979). Thus it is not necessary to assume that birds are no good at examining, in order to keep the noticing/examining distinction. There is good reason then to believe that vertebrate brain mechanisms of vision display ‘conservation of function’ rather than ‘take-over of function’ (see Chapter 5). All vertebrates have a midbrain mapping of the visual field, a mapping which is sensitive to the movement and retinal position of simple light patterns, but also all vertebrates transmit visual information directly from the eye to the thalamus in the forebrain, where modes of perception which are more cognitive than the reflexes of the midbrain may be initiated. Using this division we should expect that vision in species where a high proportion of the optic nerve is devoted to the midbrain (tectal) projections should be largely a matter of instinctive motor responses to visual stimuli which fit inbuilt feature detectors, but since all species have some forebrain involvement, perceptual learning via this forebrain pathway can never be ruled out. Also, some of the interesting achievements in vision, such as the categorisation of optical images independently of their size and position on the retina, occur in vertebrates which apparently rely mainly on the midbrain
In mammals, where the projection from eye-to-thalamus-to-striate cortex takes a higher proportion of the optic nerve, we should expect visual perception to take on a distinctly more cognitive, and less reflexive, character. It is quite possible, however, that the striate cortex projection in mammals is in some ways equivalent to the optic tectum projection in birds, even though they happen to be in different locations. Székely (1973, p. 20) points out that ‘The general arrangement of neurons in the tectum, especially their interconnections, strongly resembles that of the cortex of the higher vertebrates.’ The layering of the surface of the tectum, especially in birds, is extremely reminiscent of the layering of the visual cortex on the surface of the cerebrum, and the ‘tectal columns’ of neurons connected through the layers may serve the same function as the more celebrated cortical columns of mammals. My own view is that while this may be true in terms of local coding of the visual array, what is important for the cognitive aspects of visual perception is the interchange of successive transformations of the information originally coded in relatively faithful mappings of the retinal image. In mammals, and particularly in primates, facilities for the progressive separation and re-assembly of features derived from the retinal images are anatomically obvious in the physical adjacencies of the striate cortex, the areas immediately surrounding it, and the further reachings of visually receptive mappings into the temporal and parietal lobes to make contact with transformations of the auditory and tactile projections. The projections of the visual field to the striate cortex, the body surface to sensory motor cortex, and the basilar membrane of the ear to the temporal lobe, are found spread out over the surface of the cerebral hemispheres only in mammals. It is reasonable to expect, therefore, that the type of perceptual organisation which arises from these highly interconnected representations of sensory data on the cortex of mammals should be lacking in other vertebrate classes. But some of the differences in anatomical layout, especially those between birds and mammals, may be partly due to historical accidents. In birds, and to a lesser extent in lower vertebrates as well, sensory information from all available modalities reaches the forebrain, but only after considerable filtering and adjustment along the way. As the distinctive thing about the forebrain is that it should treat sensory data in a more abstract, flexible and selective way than is possible in lower centres, closer to the sense organs themselves, this is understandable. The mammalian forebrain is peculiar in having the maximum possible remoteness from immediate sensory input, along with pathways that seem designed to bring sensory images to it very directly. Perhaps it is this combination which brings about the controlled and detached assessment of the perceivable world which characterises human subjective experience.

If this is the case, then limitations of the perceptual pathways in the brains of non-mammalian vertebrates may be in confining selectivity and flexibility to more abstract and less detailed sensory
content. This brings us back to the ‘noticing’ and ‘examining’ division of modes of perception, with the corollary that non-mammals may be able to notice, and to react to such details as are provided by their sense organs, and also examine relatively abstract features, in the sense that they may learn to pay attention to modalities, irregularities within modalities or such complex combinations of features as are required for recognition of classes of objects. The advantages of the mammalian perceptual systems ought then to lie in the cognitive treatment of exact sensory representations. Mammals ought to be able to switch in and out analysers for many features of sensory information less amenable to optional separation in non-mammals. If the detachment of perception from immediate response is one of the advantages of forebrain mechanisms, mammals ought to be able selectively to retain small parts of sensory experiences for future use: mammals should remember more details than non-mammals. Unfortunately there is very little evidence from behavioural testing which convincingly demonstrates superior mammalian capacities along these lines, but the separation of memory of perceived events from fine sensory discriminations which may be ‘noticed’ only in terms of reflexive response is a step in the right direction.

282

The comparative anatomy of vision—conclusions

Human sight requires the receipt of an optical image by the eye and the transmission of this image, coded as the electrical firing of nerve cells, to the striate cortex of the brain, which is designed to break the code, and reconstitute the optical image in a form in which the firing of neurons indicates brightness changes, angles and movements in particular places in the scene presented to the eyes. The conscious experience of seeing requires the integrity of both the image on the retina, and the coded image in the striate cortex of the brain—damage to local regions of the retina or of the striate cortex makes people unaware of the presence of lights or objects in corresponding places in the visual field.

Other mammals utilise a similar anatomical system, with similar neural codes, and other primates, such as the rhesus monkey and chimpanzee, have relationships between eye and brain which are to all intents and purposes identical to those in man. For mammals, and especially primates, the physical apparatus available for visual perception is not radically different from that available to humans. To the extent that the subjective and cognitive aspects of human visual perception reflect the activities of eye and brain, it can therefore be argued that we share them with other mammals.

However, the striate cortex is not the only part of the brain involved in vision. If it is damaged, even in man or monkey, visual perception is radically impaired, but residual capacities to detect visual cues remain. The phenomenon of ‘blind-sight’ (Weiskrantz, 1977, 1980) demonstrates that people with lesions can correctly guess at the angle of a line which they are unable to describe in the normal way. Monkeys with extensive damage to the striate cortex can, with retraining, use information from the eyes to detect and manipulate objects (Humphrey, 1974). Other mammals are if
anything even more resistant to the effects of striate cortex loss. One of the reasons for this is that output from the eye is sent in mammals not only to striate cortex, but also to a second mapping of the visual array, in the midbrain. This midbrain centre is the main mapping apparent in birds and lower vertebrates.

An apparently reasonable conclusion is that the especially human aspects of visual perception, the detailed examination and conscious experience of what is seen, are based on the workings of striate cortex, non-mammals making do with reflexive and instinctive reactions even to the very detailed and accurate distinctions available to their excellent eyes. Of all the vertebrates only the species closely related to ourselves, the monkeys and apes, have both an eye equipped for detailed, concentrated colour vision (at the fovea of the retina) and the striate cortex brain receiving apparatus like our own. Thus one might suspect that only these primates have human-like capacities for acquiring sophisticated knowledge of the nature of objects and the relationships between them through visual perception.

Though it is reasonable, there are objections to this conclusion. The visual projection to the striate cortex has its limitations, since it could not by itself be of much help in recognising images of differing sizes, or at different positions in the visual field. And it is just this sort of abstraction which we would expect to be necessary even in lower vertebrates, such as reptiles. I have quoted evidence from behavioural experiments which suggests that species without the advantages of a visual cortex—as it happens, domesticated animals such as the goldfish and pigeon—have very considerable capacities for abstraction beyond exact retinal images and for selective combination of visual features into object categories. Therefore, although only the primates may have visual perception precisely like our own, the behavioural capacities of other kinds of vertebrate support the theory that feature analysis and combination, and the interpretation of visual patterns in terms of complex inner descriptions, are rather general characteristics of vertebrate visual systems.

A closer look at the paths travelled by visual information in nonmammalian brains suggests that the complexity of non-mammalian brain anatomy is quite sufficient for the complexity of this theory. Even if the visual projection to the more primitive midbrain centres is regarded as always more reflexive, the presence of, for instance, eleven distinct forebrain regions receiving optic nerve fibres in some species of bony fish (Réperant and Lemire, 1976), six distinct thalamic nuclei in the visual pathways of a lizard (Butler and Northcutt, 1978), and the clear post-thalamic projections of the visual pathways in birds (Nauta and Karten, 1970; Karten, 1979), show that the difference between mammals and non-mammals in the involvement of the cerebral hemispheres in the visual pathways is a matter of degree. There may well be advantages in having a complete topographic mapping of the visual field present in the hemispheres, as well as in the midbrain, and birds at least have an area in the hemispheres where the cells respond to
retinal stimulation like the cells in the striate cortex (Revzin, 1969; Karten, 1979). But projections which are not faithful repetitions of retinal patterns, but responsive rather to features of the patterns in a conditional way, depending on the experience and motivation of the animal, would be of more help in accounting for the behavioural data. Thus the extra-striate regions in mammals, and the analogous regions in birds and lower vertebrates, which interpret and direct the flow of information from the more immediate visual projections that are closer to the optical image, have considerable theoretical interest.

**Modes of perception—conclusions**

The theoretical analysis of perception implies that seeing a tree as a tree is a remarkable cognitive achievement. It is difficult to separate our own perceptions of this kind from verbal relationships in language, but if a bird sees a tree as something which can be flown around and possibly landed on and nested in, the argument is that this requires a perceptual schema of some complexity, which may be closer to an expressible verbal concept than to a succession of instinctive responses to exactly pre-programmed sensations. It is possible that much of the perception of animals is in fact composed of reflexive reactions to simple forms of stimulation, but experimental evidence of the flexibility of discriminations made under laboratory conditions implies that some of the mechanisms available to many humble creatures are of the sort which we would expect to find if sensations are taken to signify objects and arouse processes of categorisation.

Even if, in this sense, animals perceive objects, it is of course a rather limited form of cognition. But attempts to construct machines which analyse inputs from television cameras so as to recognise objects, and spatial relationships between them, emphasise that object recognition is not something which should be taken for granted (Sutherland, 1973).

**Perceiving facts and discriminating stimuli**

If we see a tree, and a squirrel climbing it, we do not merely activate inner descriptions of trees and squirrels. We can say afterwards that the squirrel was climbing the tree, that the squirrel was red and the tree was green, and that the squirrel climbed very fast, if that was the case.

We perceive not only the prerequisites for nouns but also those for verbs and adjectives, and everything else we can put into words, and we retain this knowledge, if not indefinitely, for a few minutes at the very least. Since animals lack language, it is extremely difficult to know if they perceive facts, other than the presence or absence of various objects or object categories. But there are at least two ways
in which it is obvious that the perceptions of animals are more long lasting than merely momentary detections and categorisations. The first is imitation, or mimicry. Locke suggested that the efforts of parrots to reproduce heard sounds indicates the retention of the sound perceived, as a form of memory. The instigation of action which corresponds to the initial perception demonstrates that a description or schema of the perceptual input is retained, and that it is retained in such a form that it can be translated into appropriate movements. The fact that self-observation is relatively straightforward for vocal production means that birds might simply modify their own output until it matched their auditory memory, with little need for such translation. One can perceive one’s own vocal output in more or less the same way as one perceives someone else’s. But, in the absence of video-recordings, we cannot watch ourselves doing something in the same way that we watch someone else doing it. Therefore the characteristic visually based imitation of apes implies a great deal about the conceptual organisation behind their immediate perception. A chimpanzee which puts a man’s hat on its own head may know little of the function and social significance of hats, but has perceived visually that the objects which are hats are put on the parts of the body that are heads. After watching a man put a hat on, the retinal images have to be translated into an abstract code, so that a chimp puts a hat on its own head (which it cannot see) in response to the events which it saw previously.

Animals other than apes and talking parrots do not so obviously imitate human behaviours, and when an animal imitates a member of its own species (as parrots and apes do in the absence of humans) it is less easy to distinguish imitation based on perceived facts from simpler triggering of natural behaviours. Chickens will eat if others do so, but it is not clear whether a chicken perceives ‘eating’ as an activity which either other chickens, or itself, might perform, or whether the presence of other birds moving about serves as a ‘releasing stimulus’, reflexively producing the motor responses which we call eating. Many social species of animal co-ordinate their individual activities, but there is no doubt that special purpose instincts are one of the mechanisms used.

Lorenz (1952), for instance, suggests that the sight of a conspecific flying overhead may reflexively elicit following responses in social birds. True copying of seen arbitrary body movements or co-ordinated actions may be confined to apes and man. But whenever it occurs, it demonstrates a high degree of conceptual organisation.

Apart from imitation, perceptual knowledge over and above the categorisation and recognition of sensory patterns can be demonstrated by extrapolation and inference based on perception. A dog retrieving a thrown stick is an example which should not be discounted because of its familiarity: chasers of moving prey need to be able to cut corners and run to where the quarry is going to be when they get there, not towards where moving prey happens to be when it is first perceived. More generally, the function of perception
is to direct actions, and actions may need to be determined not only 
by present sensations but also by prior perceptual experience and 
inferences based on it. The paradigm of reflexive perception is a 
fixed motor response to an unvarying input to the sense organs, but 
in human perception information is detected, stored, and made 
available for use in the future. The theme of this chapter has been 
that animal perception is not always reflexive; the internal analysis 
of current activity in the sensory nerves is usually flexible and 
conditional, showing selective attention, and may be entered into in 
brain networks which function as inner descriptions of objects or 
perceptual categories. The existence of perceptual categories may be 
manifested by the immediate response to current forms of sensation. 
Whether or not representations of perceived events are retained or 
reinstated for future use is a rather different question, one asked in 
more detail in the next chapter.