7 Reward and punishment

‘What is painful is avoided and what is pleasant is pursued’.

Aristotle, De Motu Animalium

Appetitive and aversive motivation

Up until this point, I have discussed instrumental learning and voluntary behaviour almost entirely in terms of the pursuit of the pleasant, or more automatic versions of this concept which correspond to the seeking of positive goals. This has been in the of simplicity, but clearly any theory of willed action would have to include the avoidance of undesired outcomes as well as the search for wanted rewards, and any more general and less cognitive account of the effects of motivation on learning needs equally to consider at least two kinds of motivation, associated with sought-after and feared events, or with more reflexive forms of approach and avoidance behaviours.

Perhaps even more than in other chapters, the reader may here encounter confusion due to arbitrarily selected technical terms in learning theory. The most conventional distinction between putatively pleasant versus disagreeable events refers to ‘appetitive’ versus ‘aversive’ reinforcers, corresponding to appetites and aversions or appetitive and aversive motivation. These should be regarded as the most conventionally correct terms (Mackintosh, 1974, 1983), but there are many variations in usage (e.g. Gray, 1975). I shall speak fairly loosely of attractive and aversive events, or attractive and aversive emotional states, and hope that the meaning is clear from the context. However some of the terminological difficulties arise from genuine theoretical questions surrounding the degree of interchangeability of reward and punishment. It is logically possible to conceive of a single urge underlying them both; Hull (1943) for instance based his theory on the universal biological necessities of nourishing and preserving the bodily tissues, but drew an analogy between pain and hunger as the mechanisms for dealing with these needs, and was thus able to use the single concept of drive reduction for all motivation. Behaviour in Hull’s theory is always impelled by goads, either internal or external, never attracted by equivalent positive goals. The best that one can hope for in this scheme is to minimize one’s levels of irritation and distress. Few have been optimistic enough to make quite so thorough a job of the converse of Hull’s theory — the Pollyanna conviction that the motive for response is always to make things better, life consisting of degrees of happiness, with even the most unpleasant ordeals perceived in terms of how much joyful relief the future may bring. However, Herrnstein (1969), Gray (1975), Dickinson (1980) and Mackintosh (1983) have all emphasized that escape from unpleasantness can in some cases be explained in terms of future attractions, in the context of experiments in which rats perform responses which reduce the frequency of the electric shocks they would otherwise receive.

Often there are problems in tying down the subjective aspects of positive and negative emotions to measurable behaviours, or in making even
theoretical distinctions between their effects. One can imagine building a robot in which all desirable ends were represented as positive numbers, and all adverse outcomes by negative numbers: the only motivational instruction necessary for this artificial creation would be to maximize the aggregate, and any constant, positive or negative which was added to individual values would be irrelevant. A feature of this idealized system is that the rewards and punishments, or attractive and aversive reinforcers, have equal, but opposite, effects. I shall use as a theme for this chapter the question of whether, in practice, in the natural as opposed to the idealized world, reward and punishment have this sort of symmetry. To the extent that they do not, it will clearly be necessary to say in what ways the motivational systems for reward and punishment differ.

Anatomical and functional separation of attractive and aversive mechanisms

It will be as well to start with the line of evidence just appealed to in discussing types of association (chapter 6) the biological facts of brain structures and the theories of the behavioural functions which these structures serve. It has been clear ever since Papez (1937) pointed it out that the limbic system or ‘Papez circuit’ of the vertebrate forebrain, which is an interconnected network of brain parts, is the place to look for motivational mechanisms. Lesioning of different parts produces different motivational effects. Amygdala lesions make the animal tame and inappropriately relaxed; septal lesions make it jumpy and aggressive; lesions of the lateral hypothalamus and pyriform cortex make it under- or over-sexed; and lesions of the ventro-medial versus lateral hypothalamic regions make it eat too much or too little. There are no agreed interpretations of precisely what this evidence means, but in the case of motivation associated with eating, then physiological theories have it in common that they are all complicated, assuming separate mechanisms for such factors as motivational states resulting from extreme hunger, the effects of food palatability, and detailed control over when an animal starts and stops a particular meal (Green, 1987). This supports the psychological expectation that eating may occur either as a reaction to strongly unpleasant inner sensations of hunger, or in the absence of physiological need, in considered and sybaritic anticipation of taste-enjoyment, or in some combination of these.

The stronger and more direct form of evidence is however that from animals’ reactions to mild electrical stimulation of different points in the limbic system, which gives rise to the assumption that there are pleasure and pain centres in the brain (Olds, 1958). Olds (1961) re-affirmed his belief that these results require the addition of a pleasure-seeking mechanism to any drive reduction or pain-avoidance formula such as that proposed by Hull. In Olds’s words, pleasure has to be seen as ‘a different brand of stimulation’ from the mere absence of drive (1961, p. 350). On the face of it, this arises simply because there is no obvious source of a need or drive state when rats that are not deliberately deprived of anything, run to a particular place in a maze, or repeatedly press a lever, apparently because this results in their receiving electrical stimulation of the brain (via electrodes implanted
through their skulls). Deutsch and Howarth (1963) proposed an ad hoc defence of drive theory, which relied on the assumption that the same electrical stimulation both initiated the drive and reduced it, but this cannot cope with the findings that rats will run across an electrified grid to get brain stimulation (Olds, 1961), will perform more or less normally on schedules of reinforcement for bar-pressing in Skinner boxes when long intervals intervene between successive episodes of brain stimulation (Pliskoff et al., 1965; Benninger et al., 1977) without any prior priming, and also appear to be comforted by appetitive positive brain stimulation received during illness in the taste-aversion paradigm (see p. 232 below; Len and Harley, 1974).

The behavioural effects of rewarding brain stimulation thus appear to support the view that there is an attractive motivational mechanism. But few have ever doubted this; our question is to what extent the attractive mechanism is equal and opposite to the aversive one. Some asymmetries appear to be present anatomically. Olds (1961) suggests that the reward system takes up rather a large part of a rat’s brain, the punishment system much less of it: out of 200 electrodes placed in the brain at random, 35 per cent had rewarding effects on behaviour, 60 per cent had no apparent motivational effects at all, and only 5 per cent had definite punishing effects. Using standard behavioural tests, the precise location of rewarding and punishing sites can be plotted; in Olds and Olds (1963) study a point was judged attractive if animals pressed a lever to turn electricity on, but aversive if, when a train of stimulation was started by the experimenter, the rat pressed a similar lever to turn this off.

The main features of the anatomical lay-out suggested by this procedure are that:

(i) points where electrical stimulation is attractive are centred generally on the hypothalamus and its main fibre tract connection with the septal area, the median forebrain bundle, with hardly any involvement of the thalamus (a sensory relay station);

(ii) conversely points with exclusively aversive behavioural effects were found frequently in the thalamus, and also in the periventricular region of the midbrain

(iii) many points which showed both attractive and aversive effects were found in the hypothalamus, in the medial area for instance;

(iv) ‘pure’ effects one way or the other were most likely in fibre bundles, while the ambivalent points, in nuclei, demonstrate that the two systems are often brought close together in physical proximity.

These physiological results do not provide strong evidence as to whether punishment is the mirror image of reward in terms of its behavioural effects, but they certainly suggest that there are two separate physiological systems, which interact, and that the aversive system is fairly directly connected to sensory input, in the thalamus and midbrain, as would be expected for pain and discomfort, whereas the attractive mechanism is intimately involved with metabolic and autonomic control, as would need to be the case if some of the attractive systems serve purposes in connection with bodily needs and homeostatic balances, and cyclical variations in behaviour. This can be related to the analysis of different types of drives (Gray, 1975) and to theoretical schemes of biological function. At a very rudimentary stage of examination of this, it would not surprise us to find that there were motivational imperatives.
of different degrees of urgency. A hungry animal being chased by a predator should only have one choice when internal comparisons are made between the importance of eating and the importance of escaping, but, while keeping a watchful eye open to the possibility of danger, a prey animal may need to make sophisticated adjustments about its own choices of palatable but costly versus abundant but boring items. In terms of function, it seems unlikely that the underlying mechanism for panic flight should have much in common with the incentive to fill oneself with the most energy-rich food available in times of great abundance. And in the natural world, as opposed to the laboratory, for many species the time devoted to active escape from danger or the immediate food-seeking may be short by comparison with that taken up by nest construction, complicated social interactions of several kinds, migration, and exploring and updating of unfamiliar or familiar territorial domains. All these various activities need some kind of psychological system to sustain them, and it is not likely that just one or even just two kinds of motivational apparatus would be sufficient for the whole lot.

Similarities between reward and punishment

Having established that there are grounds for expecting qualitative differences between attractive and aversive motivational systems, we ought now to inspect the contrary evidence — that the behavioural effects of the two systems are roughly equal but opposite. That is to say, attractive stimuli attract, and thus encourage the performance of responses which it has been learned will bring them about, while aversive stimuli repel, and discourage behaviours which make them more likely. The above statements may appear to be tautologous, and thus not worth experimental examination. This is almost the case, and perhaps would have been had not both Thorndike (1931) and Skinner (1953) argued the contrary. Thorndike was persuaded by some data that should have been treated more tentatively that neither young chicks nor undergraduates possessed any mechanism which would prevent them from doing again things which had previously proved disadvantageous, and from the beginning had emphasized that it was accidental successes, rather than accidental error, that was the engine of trial-and-error learning. Skinner was similarly sceptical about the ability of rats to associate unfavourable outcomes with their own behaviour, but this was linked to an idealistic and perhaps practically sound rejection of the use of punishment by parents and teachers to control the behaviour of children. Skinner’s argument seems to have been that punishing a child for wrong-doing will produce generally counter-productive emotional upheavals, which may become transferred even more counter-productively to associated events by classical conditioning; but that the punishment will not act as a deterrent for any specific response.

We may accept Thorndike’s suspicions about the fallibility of chicks and undergraduates, and Skinner’s doubts about the advisability of punitiveness in parents and teachers, without discounting the symmetry that certainly exists to some degree between the encouragement of responses by reward and their deterrence by punishment, but without perhaps going quite as far as to say that
‘The most important fact about punishment is that its effects are the same as those of reward but with the sign reversed’ (Mackintosh, 1983, p. 125). The deterrent effect of aversive stimuli on instrumental responses can in fact be readily demonstrated in the typical Skinner box, and indeed was so demonstrated by Estes (1944). If rats press a lever because this delivers food pellets, they may be deterred from pressing it by the addition of mild electric shocks, delivered to the feet at the moment the lever is pressed. Depending on their degree of hunger, the size of the food pellets and the strength of the shock, they will continue to press if rewarded with little or no punishment, and continue to refrain from pressing if the punishment is strong enough, and is given invariably if they occasionally try to get away with it. Moreover, if the rewards cease, the effect of the rewards will dissipate as the animals learn that the response no longer brings them about, and similarly if shocks cease, the effect of punishment will disappear as the animals learn that these are not forthcoming: these effects are symmetrical, since they can both be construed as learning about the consequences of responding (Mackintosh, 1974, 1983).

However, there are limits to the symmetry. First there is a logical difference between learning about positive and negative response consequences. Since positive consequences are sought, and, if learning has been successful, found, then any change in the positive consequences of responding will quickly become apparent to the responding animal. On the other hand, since negative consequences are avoided, then changes in negative consequences may not immediately present themselves to the animal which is not responding, and this is one reason why, other things being equal, we might expect the deterring effects of a temporary unpleasant consequence of responding to be somewhat more lasting than the encouraging effects of a temporary incentive with equivalent emotional force. This may be the explanation for the finding of Boe and Church (1967) that a very strong series of shocks given consistently for rats’ intermittently food-rewarded lever pressing deterred further lever-pressing completely and indefinitely.

It is certainly arguable that in addition to bias against gathering new information about pain and distress, any scale of these affective qualities will be difficult to map on to a scale of the desirability of food pellets according to their size or taste, even with a change of sign. However, within limits, it is possible in behavioural experiments to construct a scale of practical equivalences, by setting off given amounts of attractiveness in a goal against degrees of unpleasantness encountered in the course of achieving it. Vast amounts of evidence were collected before the Second World War by Warner (1927, 1928a, 1928b) and Warden (1931), among others, using the ‘Columbia obstruction box’, in which rats were required to run across an electrified grid to obtain access to food, water or a member of the opposite sex, under systematically varied conditions. Rats were reluctant to run across the standard grid for food until they had been without food for at least two or three days, but crossed with little hesitation for water if deprived of this for 24 hours. Male rats ran across the same grid to get to a female in heat rather more often one day after previous sexual contact than four weeks after, and very much less if tested within six hours of previous copulation; females only crossed at all to males in the most receptive half of their estrus cycle, with a peak number
of crossings confined to the estrus phase. The highest rates of crossing the standard electrified grid were observed in maternal rats separated from their young (Nissen, 1930; Warden, 1931).

It would be unwise to place very much emphasis on these results, but it is clear that the animals were capable (a) of learning that there was a desired goal object of a certain kind on the ‘incentive compartment’ on the other side of the electrified grid, and (b) of combining this knowledge, in however simple a form, with the level of their current appetite, so that, for instance, they would cross to food livery hungry, but not when moderately so. Stone (1942) was able to get essentially similar results, without using the somewhat artificial device of the shocking grid, by training rats to dig through tubes filled with sand, or to scratch their way through a succession of paper-towel barriers blocking a runway, in order to get to goal objects. More precise quantification was obtained by Brown (1942) and Miller et al., (1943: see Miller, 1944), who trained rats to run down an alley towards food while wearing a harness with a cord attached, which allowed their movements to be carefully measured, in some cases in terms of the force with which they pulled against a calibrated spring. It was found that hungry rats, trained to run towards food, pulled against the spring with almost as much force when 2 metres away as when very much closer to the food. However other rats, who received electric shocks in the goal box instead of food, pulled vigorously to get away when subsequently placed close to the goal box, but did not pull at all when placed 2 metres away, even after extremely severe previous shocks. It would have been odd if any other result had been obtained, since hungry rats are presumably still hungry when far away from food, whereas shocked rats are not necessarily afraid once they are far away from the site of their aversive experience (see below, p. 218). The difference between the pulling towards food and away from shock is often referred to as the difference between approach and avoidance gradients, and drawn as in Figure 7.1. The argument in favour of such approach-avoidance gradients is strengthened by experiments in which the same rats are both shocked and fed in the same goal box. Subsequent behaviour at various points on the path to this goal can be predicted in terms of the strength of current hunger, and the intensity of previous shocks. With moderate values of both, animals approach about half-way towards the goal and then stop, as would be expected from Figure 7.1. Either stronger hunger or weaker shock leads to closer approach; with weak hunger or more aversive shocks animals naturally keep further away. Although this was true on average, there was considerable variation between and within individual animals. Some rats adopted a pattern of consistent vacillation, of increasingly hesitant approaches followed by abrupt retreats, while others moved forward in steps, making long pauses before each small approach, eventually coming to a complete halt (Miller, 1944).
Figure 7.1 Approach and avoidance gradients.
Schematic plots of how the strength of approach and avoidance responses may vary with distance from the goal object based on experiments in which rats receive food and electric shock at the same place. In (a), it is apparent that strong approach tendencies may result in high points of avoidance gradients being encountered. Paradoxically, (b) demonstrates that a reduction of avoidance tendencies in circumstances of conflict may have the effect of raising the point at which approach and avoidance tendencies balance out. After Miller (1944).

All this suggests that positive incentive, or the attractive-ness of a goal, can be somehow weighed against the negative incentive derived from previous aversive experiences. Logan (1969) used precisely these terms in another claim that 'the effects of punishment are symmetrically opposite to the effects of reward', based on an experimental variation on the theme of conflict between reward and punishment, which included a more explicit choice between alternatives. Rats were allowed to choose between running down a black or a white alley, after having previously done ‘forced trials’ to ensure equal experience of what the black and white choices entailed. First, preferences were established by such differentials as seven food pellets at the end of the black alley, but only one in the white; or three pellets given at the end of both alleys, but available immediately in the white goal box, but only
dropped in 12 seconds after arrival in the black goal box. Both these procedures establish strong preferences in hungry rats, since these behave as if they would rather have seven than one food pellet, and a given amount of food sooner rather than later. Logan then examined how easy it was to reverse these preferences by the obstruction box method of making the rats run over an electrified grid for 2 feet before reaching the preferred goal, and varying the intensity of the shocks thus delivered. A very orderly effect of shock intensity on percentage of choices of the originally preferred goal was observed, and a stronger shock was necessary to persuade the animals to choose one instead of seven pellets than to shift the preference for immediate versus delay rewards of the same size. This difference was even more pronounced when shock had to be endured on only 50 per cent of the approaches to the preferred goal. Choice of seven versus one pellet was very resistant to that procedure, the risk of only a very strong shock reducing preference to just under 50 per cent, whereas the choice of immediate over delay reward was still strongly determined by shock intensity, rats settling for delayed rewards fairly frequently (on about 60 per cent of choices) even with the risk of only a low-shock intensity (Logan, 1969, p. 47).

All these results, and many others (Azrin and Holtz, 1966; Solomon, 1964; Morse and Kelleher, 1977) seem to suggest that reward and punishment are ‘analogous if not equivalent processes’ (Morse and Kelleher, 1977), are symmetrical but opposite in their effects, and so on. Gray (1975, pp. 135 and 229) has formalized this view first with respect to the symmetry of the possible behavioural effects of delivering or withholding attractive and aversive events; and second by presenting a theoretical model, shown in Figure 7.2, in which, as can be seen at a glance, precisely comparable mechanisms are proposed for the operation of reward and punishment, with a ‘decision mechanism’ which allows for the results quoted above, in which the attractive effects of reward are balanced against the aversive effects of punishment. In instrumental learning therefore, there are many reasons for assuming that punishment may sometimes operate in more or less the same way as reward, even though there are differences in anatomical factors and in ecological function. In terms of Figure 5.8, which was used to summarize the sorts of associations possible in instrumental learning with reward, all that is necessary is to substitute ‘unwanted’ for ‘wanted’; to interpret ‘appropriately associated with unwanted events’ to mean that such behaviours will involve withdrawal rather than approach; and thus to suppose that the result of learning that a response has unwanted consequences (at (1) in Figure 5.8) will be an impulse to inhibit such responses rather than make them. As in Gray’s model (Figure 7.2), it is necessary that rewards are automatically linked to impulses to approach, and to the repetition of rewarded responses, while aversive stimuli must be inherently linked to withdrawal, behavioural inhibition, or an internal ‘stop’ command. Clearly, as a consequence of this in general and in Figure 5.8, when the punishment mechanism works, punished responses are suppressed, and the relevant motivating event is notable by its absence).
The theory of avoidance learning

The symmetry of attractive and aversive events can certainly be maintained in plotting the predicted effects of increases and decreases in their frequency. Animals should behave so as to maximize their receipt of appealing experiences and minimize their encounters with aggravation and distress: thus they should learn to repeat responses which either bring about or prolong rewards or prevent or cut short punishments; and they should also learn to inhibit responses which either prevent or truncate pleasurable or satisfying states of affairs, and they should learn to inhibit responses which initiate or continue pain or distress. Both this last sentence and Gray’s diagram (Gray, 1975, p. 135) may appear complex, but they are simply behavioural elaborations of the pleasure/pain principle. The first step in this is to say that responses which bring about rewards should be repeated, and responses that bring about punishments should be stopped. The second step goes beyond this, to deduce that responses which prevent otherwise available rewards should be inhibited, whereas responses which prevent or truncate otherwise imposed punishments.
should be repeated. It has often been pointed out (Mowrer, 1939; Dickinson, 1980) that this second step is very much more demanding of the cognitive abilities of both the animal and the learning theorist, because the critical consequences of responding are unobservable — what is important is that nothing happens.

There are a number of explanations for why the absence of an event may be critical in serving as a goal or reinforcement for instrumental learning. Perhaps the most straightforward explanation for the theorist, if not for the system being explained, takes the form of assuming that the behaving system contains comparator mechanisms, which assess whether current levels of attractive or aversive stimulation are greater or less than expected. If an expected event does not take place, this fact can thus be fed into the relevant motivational device — the absence of a reward should be regarded with displeasure, but the absence of an expected punishment clocked up as something to be sought after. Such arrangements are included in Figure 7.2. The main problem with this is that it takes an enormous amount of continuous cognitive processing for granted. Whenever a normally obtained reward or punishment is missed, the system should sit up and take notice, and this implies some form of continual vigilance. But we have already seen that such comparator mechanisms, albeit of varying degrees of complexity, are a universal feature of basic learning processes. In habituation to motivationally insignificant stimuli, it is assumed that all stimulus input of this kind is compared to a ‘neuronal model’ of what is expected, the distinction between novel and familiar stimuli being between stimuli which do or do not match the model (Sokolov, 1963 see p. 40). For classical conditioning, it is assumed that the signalling stimulus arouses some representation of the signalled event, subsequently compared with obtained experience (Dickinson, 1980, see p. 105). For instrumental learning with rewards, we assume that the representations of wanted events are available, often before the relevant response is made (Tinkelpaugh, 1928, see p. 153). When an expected reward is not obtained for an instrumental response, most theories assume that some process of frustration or inhibition is aroused, which is responsible for the eventual decline of non-rewarded responses (Rosenzweig, 1943; Amsel, 1962; Pearce and Hall, 1980; Dickinson, 1980; Gray, 1975). On these grounds it would seem almost an aesthetic necessity that for the purpose of achieving harmony and symmetry, we should also assume that when an expected punishment is omitted this is sufficient to encourage the repetition of any response associated with the omission. Fortunately there is behavioural evidence to suggest that something of this sort does indeed take place (Herrnstein, 1969). However, there is an even greater amount of evidence to suggest that this is not the only significant process in instrumental learning motivated by aversive stimulation.

**Escape learning**

Little attention has been given to what is formally known as escape learning, in the case where electric shocks or other localized aversive stimuli are delivered, but most of the arguments about Thorndike’s experiments on cats which escape from small boxes would apply — for instance, is the successful response an automatic habit, or is it made in knowing anticipation
of its consequences? If a rat in a Skinner box is exposed to continuous painful electrical shocks from the floor, it will normally learn rapidly any response which makes this stimulus cease, whether it is moving to a part of the floor that is safe, rolling on its back to make use of the insulating properties of its fur, or pressing a standard lever which serves as the off-switch. It is arguable that responses made to relieve already present pain or discomfort are more likely to be made automatically and reflexively than topographically similar behaviours learned under the influence of rewards which follow them. First of all painful aversive stimuli may have greater motivational immediacy than others (see below, p. 230), but apart from that, there is little need to construct cognitive representations of a motivationally significant stimulus which is already present before the response is made, whereas the logical structure of learning for rewards means that there has to be an internal and cognitive representation of the motivating event, if the reason for the response is to be known while it is being initiated. In Hullian terms the drive stimulus may be rather more obvious and vivid when it is externally imposed than when it is generated by internal time-sensitive cycles (see Gray, 1975; chapter 4). Nevertheless, one of the behavioural phenomena reliably observed when rats press levers in Skinner boxes to turn off electric shocks is difficult to explain purely on a Thorndikean stamping-in basis. In all experiments of this type, it is necessary to specify how long the shock is turned off for. It can be for the rest of the day, or the rest of the experiment, but more commonly it is for something like 20 seconds, after which the shock starts again, and the lever must be pressed again (Dinsmoor and Hughes, 1956; Davis, 1977). Under these circumstances rats have a strong tendency to hold the lever firmly pressed down during the shock-free intervals. In a very large part, this is due to the species’ instinctive reaction of ‘freezing’ or crouching very still, which is elicited by painful stimuli or signals of danger (Davis, 1977), but it is also maintained by its utility as a preparation for rapid execution of the next response (Dinsmoor et al., 1958; Davis, 1977).

Apart from the emphasis on instinctive reactions, few conclusions can be drawn from the fact that animals learn rapidly to repeat naturally favoured responses which turn off shocks (Davis, 1977). A great deal more theoretical interest has been attracted to the case of avoidance learning, since by contrast with escape learning, where the motivating stimulus occurs conspicuously before each response, this event, when learning is successful, is rarely, if ever, seen.

The two-process theory of avoidance learning

The two-process theory of avoidance learning appeals to the classical conditioning of fear or anxiety as the first process, and the instrumental reduction of this unpleasant emotional state as the second. It is associated with the names of Mowrer (1940, 1960) and Miller (1948), and an apparatus known as the Miller-Mowrer shuttle box (actually used originally by Dunlap et al., 1931). This is a box with two compartments, each with a floor through which electric shocks can be delivered, perhaps with a hurdle or barrier between them. Every so often, say about once every 2 minutes, a buzzer is turned on
for 10 seconds. If the animal in the box stays in the compartment where it is when the buzzer starts, it receives shock at the end of the 10 seconds. However, if it jumps or otherwise shuttles to the alternative compartment before the 10 seconds are up, the buzzer is (in most experiments) turned off, and (in all experiments) no shocks are given on that trial. It is clearly greatly to the advantage of the animal concerned if it shuttles between the two compartments when it hears the buzzer, and thus the basic result is capable of explanation by the principle that behaviours which reduce the frequency of unpleasant experiences should be learned (Herrnstein, 1969).

Thus, instrumental learning, conceived as the principle of reward and punishment in an abstract logic of events, is capable by itself of explaining why avoidance learning ought to occur. It cannot explain why, in many instances, mainly when the required responses conflict with instinctive reactions, avoidance learning fails to occur (Bolles, 1978; Seligman and Johnston, 1973) and it misses out altogether the undeniable fact that in most cases the delivery of aversive stimuli arouses distinctive emotional states, which are highly conditionable, in the sense that predictable though not identical emotional reactions are quickly induced for other stimuli which are taken as signals for impending pain or distress (see backward conditioning, p. 87). There is thus ample reason to retain the two-process theory, to the extent that it predicts both conditioned emotional states and responses motivated by them, while also acknowledging that there is evidence for more calculated forms of avoidance learning, based on anticipation of the consequences of responding, compared to the state of affairs that might otherwise be expected to obtain (Gray, 1975; Mackintosh, 1983).

The paper ‘Anxiety-reduction and learning’ by O.H. Mowrer (1940) stated a simple and direct form of the two-process theory, including the assumption that anxiety reduction should qualify as a ‘satisfying state of affairs’ in Thorndike’s Law of Effect. Behavioural responses which bring relief from anxiety should therefore be stamped in or fixated. As an experimental test of the theory, Mowrer trained several groups of rats and guinea pigs to run around a circular track composed of eight grid segments which could be independently electrified. Once a minute a tone was turned on for 5 seconds, with shock to be delivered to the segment the animal was initially standing on at the end of the tone. But as soon as the animal moved forward at least one segment the tone was turned off, and no shock was delivered. After three days of this at 24 minutes per day, the animals received only two out of the 24 potential shocks in a day, and stable performance was reached at four shocks per day which counts as about 80 per cent (20/24) correct avoidance. There were minor species differences between the rats and guinea pigs, rats learning better with random rather than fixed intervals between trials, and the guinea pigs the other way around. Mowrer (1940) argued that the running response to the tone was established because it relieved conditioned anxiety or dread, but pointed out that learning of this sort appeared to work best when the response that ended dread closely resembled the response which would be used to escape from the dreaded event itself, when it was present. This would be reasonable, even if the main process involved was the conditioning of an emotionally loaded representation of the stimulus event, but it leaves Mowrer’s evidence open to the objection that no
conditioned anxiety, or anxiety relief, was necessary, since more direct ‘knee-jerk’ conditioning of the motor response alone would account for the results.

There are many reasons why this objection could be discounted, but the clearest demonstration that avoidance learning involves more than motor response shift would require that the response made to avoid is quite different from that used in escaping from the event avoided. Such a demonstration was provided by Miller (1948), who used a procedure in which the response learned was separated from the response elicited by shock, both by topography and the lapse of time. In a two-compartment shuttle box Miller first gave 10 trials only in which rats received intermittent shock for 60 seconds in a white compartment without being able to escape, and were then confronted with continuous shock and a suddenly opened door. All the animals here learned to run through the door quickly, to the safe black compartment. Also when subsequently put in the white compartment even without shock, they ran to the black compartment. The next phases of the experiment showed that this was not just a matter of automatic running. First, the rats were left in the white compartment with the door closed, but a wheel by the door which, if turned, opened the door. All rats showed variable behaviour around the door which could be construed as attempts to get through it. Half of them (13/25) moved the wheel by accident the fraction of a turn necessary to open the door, during the first few trials, and thereafter became more and more adept at turning the wheel to open the door as soon as they were placed in the white compartment. The others tended more and more to adopt the posture of rigid crouching. These results are strong evidence that the initial phase of shocks meant that the rats thereafter were put into an aversive motivational state by being in the white compartment, and that the novel behaviour of turning a paddle-wheel device was learned by being associated with escape from the anxiety-provoking compartment. Miller (1948) then changed the procedure for the 13 rats that were turning the wheel, so that the wheel no longer worked, but the door would open if a bar, on the other side of it, was depressed. The first time this was done, the animals turned the wheel more vigorously than usual — although only a fractional turn had previously opened the door the typical (median) number of whole turns was almost 5, one rat making 530 whole turns. However, all but this one, by the fifth attempt, quickly opened the door by pressing the bar, instead of turning the wheel.

**Problems for two-process theory of avoidance**

The experiment by Miller (1948) and others in which animals learn to shuttle back and forth between two compartments as the signal for impending shock (Mowrer and Lamoreaux, 1946; Kamin, 1956) appear to support the two-process explanation of classical conditioning of what can loosely be called fear or anxiety to external cues, and instrumental reinforcement of responses which remove the animal from these cues, or otherwise reduce fear. The reader may however have noticed that there is a blatant contradiction between what is implied in this account and the claim made in the previous chapter (pp. 184—8) that classical conditioning is ineffective if there is only intermittent pairing of a
signal with a signalled event. It is in the nature of successful avoidance learning that the signal for shock is no longer a signal for shock, because responses are made which prevent the shock happening. There are a number of ways around this contradiction. Most directly, there is 100 per cent correlation between the signal and the shock when no response is made, and the fact that a response has been made can quite properly be regarded as having altered the character of the signal — ‘signal alone means shock’, and ‘signal plus response means no shock’ are both reliable and consistent associations. There are other aspects of reactions to aversive stimulation however which indicate the need to appeal to special factors to do with strong instinctive reactions to pain and danger, and with related emotional and physiological changes involving stress.

The contradiction between the supposed sensitivity of classical conditioning to the degree of correlation between events, and the persistence of avoiding behaviours in the absence of anything to be avoided, is taken to extremes in what is called ‘traumatic avoidance learning’ (Solomon and Wynne, 1954; Solomon et al., 1953; Solomon and Wynne, 1953). This is simply a shuttle-box avoidance procedure used with dogs rather than rats, and with strong shocks. With a 3—minute interval between trials, and a 10—second signal of lights going off and the gate between compartments being raised, Solomon and Wynne (1953) reported that dogs received only seven or eight shocks before reaching the criterion of 10 consecutive avoidance responses, which involved jumping over a shoulder-high hurdle. This is not particularly exceptional — the result that is theoretically important is that once this criterion had been reached, dogs showed no sign at all fever reducing their tendency to jump when the signal came on, even after 200 trials at 10 per day, or in one case after 490 trials. There was evidence that the animals became much less emotional as the extinction procedure (without any received or potential shocks) proceeded, but that the latency of jumping either stayed the same or decreased.

One argument is that these two reactions are connected, and that ‘anxiety conservation’ occurs because the continued fast jumping prevents the dog ‘testing reality’ by discovering that shocks will no longer occur if the signal is ignored. However, it was no simple matter to confront the dogs with reality in a way which quickly removed the conditioned response of jumping. If a glass barrier was inserted between the compartments so that successful jumping was impossible, the dogs at first appeared excited and anxious, and over 10 days quietened down, but then, if the glass barrier was removed, they immediately began jumping again. An alternative procedure of punishing the jumping response was tried, by simply arranging that shock was present in the opposite compartment to the one the dog was in, when the signal sounded. Here it was highly disadvantageous for the dogs to continue jumping, but this is precisely what most of them did. If the glass-barrier procedure and the punishment procedure were alternated, then the jumping response was finally suppressed, but its persistence in the absence of any overt benefit was clearly not predictable on the straightforward version of the two-process theory (Solomon et al., 1953; Solomon and Wynne, 1954; Seligman and Johnston, 1973).

It is therefore necessary to modify the two-process theory in some way to account for the persistence of avoidance responding when no further
motivating events are observed. There are several possible modifications, all of which have some merit. However, it first has to be said that modification of the two-process theory is not always necessary, since avoiding behaviours are not always very persistent. For instance in Mowrer’s experiment already quoted (1940), in which rats or guinea pigs ran around a circular maze at the sound of a 5—second tone, omitting all shocks from the procedure meant that many animals stopped running by the end of one session of 24 tones. Since they had previously only been receiving about four shocks per day, this implies that

the first few times they waited the full 5 seconds of the tone and received no shock, this immediately led to a drop in the tendency to run. Many other experiments with rats have found conditioned avoidance responses quickly declining when shocks are no longer given (Mackintosh, 1974, 1983). In most of these the shock is not associated with the animal being in a particular location (as it was in the experiment of Miller, 1948) but with a light or sound signal. Thus the only modification in two-process theory that is necessary is the one which suggests that the signal with no response functions as the cue which is consistently associated with shock, while the signal together with a response is associated with not getting shock (Dickinson, 1980). It is already the case of course, that the instrumental part of two-process theory should be strengthened by intermittent reinforcement (see p. 183).

**Habitual responding which prevents fear**

As Mowrer’s original formulation of two-process avoidance learning (1939) was explicitly inspired by Freud’s book *The Problem of Anxiety* (1936), the next modification is no novelty. Freud of course proposed that although anxiety was a reaction to danger, the important dangers for people were internal personal conflicts, rather than external events, but whatever the source of anxiety, it is a thoroughly unpleasant and unwanted emotional state. One of Freud’s main points was that ‘symptoms are only formed in order to avoid anxiety’ (1959, p. 144). In other word, neurotic habits keep anxiety in abeyance. The same principle appears to apply to avoidance learning. Solomon and Wynne (1954) pointed out that overt signs of emotionality (e.g. pupil dilation, excretion) in their dogs tended to decline after initial experiences, partly because the avoidance response of jumping was made so fast there was no time for the autonomic nervous system to react to the signal. But the general reactions of the animals between the signals also became much calmer. Measurement of the heart rate of dogs (Black, 1959) and of the behavioural reactions of rats (Kamin *et al.*, 1963) support the contention that with well-trained avoidance responding there is little evidence of conditioned fear (Seligman and Johnston, 1973).

Thus the two-process theory has to be extended beyond its most rudimentary formulation, which implies that fear is classically conditioning, and responses are only made when impelled by high levels of this conditioned fear. It is possible that well-learned avoidance responses may be sustained for some time purely by habit (Mackintosh, 1983, p. 168), but it is also likely that, as Freud suggested (1959, p. 162) there is a second kind of anxiety, which motivates the avoidance of full-blown fear reactions. One aspect of this ‘non-
fearful’ motivation for avoidance responses that has been put to experimental test is that responses may be made as if they were being rewarded by safety or relief, as ‘attractive events’ (Dickinson, 1980, pp. 106—9). It is certainly the case that explicit environmental signals which guarantee the absence of shock facilitate performance when given as feedback for avoidance responding by rats (Kamin, 1956; Morris, 1975). However, this is only a relative kind of attractiveness:

according to theoretical definitions, the absence of shock can only be attractive, indeed it can only be noticed, if there is already some form of anticipation of shock (Dickinson, 1980). Responding in anticipation of safety from pain can hardly be equivalent in its emotional connotations to responding for palatable titbits, as is suggested by the finding that even animals responding successfully on avoidance schedules may develop stomach ulcers (Weiss, 1968; Brady et al., 1958). However, it seems undeniable that standard avoidance learning procedures involve the avoidance of conditioned fear, as well as escape from conditioned fear. Well-trained animals do not wait until they become afraid before they respond, they respond so as to prevent themselves becoming afraid.

**Herrnstein’s theory of avoidance learning**

Herrnstein (1969) proposed that the notion of fear, or indeed of any emotional evaluation whatever, should be eliminated from theories of avoidance learning by adapting the hypothesis above so that it refers only to external aversive events; animals respond so as to prevent themselves from experiencing aversive events, or ‘the reinforcement for avoidance behaviour is a reduction in time of aversive stimulation’ (1969, p. 67), aversive stimulation being interpreted only in terms of events observable in the environment, outside the animal. Thus one of the processes in two-process theory appears to be eliminated. However, Herrnstein recognizes that, in order for his theory to work, it must be assumed that animals first detect the changes in the likelihood of being shocked with and without a response, and secondly, learn to produce activities which have the outcome of lessened exposure to disagreeable events. His explicit alternative to ordinary two-process theory is to substitute a more bloodless cognitive assessment of shock probabilities for internal emotional states which have motivational properties. This is a technical possibility, in that once it is assumed that behaviour will be directed by the outcome of less of a certain kind of experience, it is not absolutely necessary to add in anything else by way of emotion. I shall conclude that independent evidence suggests that in practice there usually are conditioned emotional effects produced by external aversive events, but it is appropriate to give first Herrnstein’s side of the argument. The usual two-process theory makes most sense when there is a clear signal which predicts an avoidable shock. Thus when the buzzer sounds in a shuttle box it appears reasonable to assume that the buzzer will at least initially arouse fear, which will motivate the learning of new responses. But it is possible to study avoidance learning by other methods, in which there is no clear signal for impending shocks. Sidman (1953) discovered a procedure known as ‘free operant’ or Sidman avoidance, in which rats press a lever in a Skinner box to avoid shocks. There is no external signal. One timer ensures
that a brief shock will be delivered every $x$ seconds if there are no responses, and another timer over-rides this with the specification that shocks are only delivered after $y$ seconds since the last response. Thus if the rats do not press the lever they will be shocked every $x$ seconds (say 10) but if they press the lever at least once every $y$ seconds (say 20) no shocks whatever will be delivered. It is possible to adapt two-process theory by assuming here that there is an internal timing device which serves as a signal — the sense that $(y-1)$ seconds have elapsed since the last response could serve as a signal for impending shocks, and there is some indication that this sort of thing happens, since rats often wait for $(y-1)$ seconds before responding. However, Herrnstein and Hineline (1966) modified this procedure to discount internal timing, by making all time intervals random. In their experiment, rats were shocked on average once every 6 seconds if they did not respond, but when they pressed a lever, they produced a shock-free period which averaged 20 seconds. They could not postpone shock indefinitely, but they could reduce the frequency of shocks significantly by lever-pressing — it is as if they could escape temporarily from trains of shocks only 6 seconds apart on average. This procedure was in fact more successful in inducing lever-pressing than Sidman avoidance (Herrnstein, 1969). Since all time intervals were probabilistic, Herrnstein argues that any internally timed process of fear would be redundant, for both the animal and the theorist, and both would do better simply to accept that lever-pressing reduces shock frequency, and is therefore a behaviour worth performing (1969, p. 59). There is something to this, as a parsimonious strategy or procedure, but there is compelling additional reason to suppose that inner emotional states may be aroused, whether or not they serve a useful function in a particular experimental procedure. Monkeys which perform for long periods on a Sidman avoidance response, but so successfully that they hardly ever receive shocks, nevertheless are liable to develop stomach ulceration which may prove fatal (Brady et al., 1958). Seligman (1968) found that random and unpredictable shocks produced ‘chronic fear’ in rats, assessed both by stomach ulceration and by very substantial depression of food-reinforced responding in the ‘conditioned emotional suppression’ technique. These are precisely the sorts of data that suggest the involvement of a central emotional state, over and above whatever cognitive assessment of shock frequency might be sufficient grounds for instrumental behaviour. It thus seems more likely than not that even Herrnstein and Hineline’s (1966) rats felt relatively relieved once they had pressed the lever, but were aversively motivated when they did not. The best tests of emotionality in these circumstances are undoubtedly physiological indices, but some aspects of the behavioural data are more suggestive of strong emotion than cognitive finesse. After experience on the schedule described above, in which lever presses were necessary to produce intervals between shocks averaging 20 seconds, the rats were left to respond with all shocks programmed at this average, irrespective of responding; in other words responding was completely without utility, and it eventually ceased. However, this extinction process was extremely prolonged, and therefore has something in common with the persistent behaviour observed in Solomon et al.’s (1953)
dogs. One rat made 20,000 responses during 170, 100-minute sessions in which responses were useless, before slowing to a halt. This may be because of the strength of an automatic habit, but it is likely that this persistence of the behaviour is related to the emotional and motivating force of the painful aversive stimuli. With a very similar procedure, in which rats had the opportunity of distinguishing between circumstances in which randomly delivered food pellets either did or did not depend on lever-pressing, no such persistence of unnecessary behaviour was observed (Hammond, 1980). This again suggests some form of asymmetry between the motivating effects of attractive and aversive events, with, if anything, aversive stimuli having more powerful and long-lasting emotional effects than attractive ones. Herrnstein (1969) was right to point out that the behavioural evidence to define emotional effects is often lacking, and his arguments support those of Freud (1936), Mowrer (1939), Solomon and Wynne (1953) and Seligman and Johnston (1973), that behaviour which is motivated by the avoidance of aversive events can often be sustained with little overt or covert sign of high emotional arousal. But this does not mean that emotional states can simply be dropped from all discussions of the motivating effects of aversive events; behavioural evidence including posture and other instinctive forms of emotional expression and results obtained with the ‘conditioned emotional response’ (CER) procedure, as well as more direct physiological indications of autonomic arousal, plus all the data not included here on the effects of tranquillizing drugs on aversively motivated performance (see e.g. Gray, 1982; Green, 1987), provide ample grounds for continuing to include conditioned fear or anxiety in theories of aversive motivation.

### Instincts and anticipation in avoidance responding

It is difficult to determine to what degree such responding, which avoids anxiety, is based on habit, as opposed to calculation of the undesirable consequences of not responding. It is probable that dogs and cats, if not rats, have a certain degree of anticipation of specific painful possibilities which they wish to avoid. Thus dogs or cats will strongly resist being put back into an apparatus in which they have been shocked several days before (Solomon et al., 1953; Wolpe, 1958). Rats do not normally do this, in fact there is often a ‘warm-up’ effect, meaning that after a 24 hour break, rats do not respond in an avoidance procedure until they have received a number of shocks. Although avoidance responding of a minimal kind, usually the continuous flexing of a leg which will be shocked if it is extended, can be obtained in cockroaches and spinal mammals (Horridge, 1962; Chopin and Buerger, 1975; Norman et al., 1977), it is worthy of note that decorticate rats, though they may be trained to perform many standard food-rewarded tests, have never been reported to have mastered any of the avoidance learning tasks used with normal animals (Oakley, 1979a). This proves nothing in itself, but adds to the general impression that avoidance learning, or perhaps anxiety itself, is partly a product of the imagination. Ordinary rats typically perform avoidance tasks at 80 per cent success, receiving one shock in five, while it is much more common for cats and dogs, and rhesus monkeys, to perform a response almost indefinitely after receiving just a few shocks (Solomon and Wynne, 1954; Wolpe, 1958; Brady et al., 1958). Much anecdotal evidence, for instance
about monkeys looking for snakes, suggests that larger mammals have specific expectancies about precisely what aversive stimulus is to be anticipated, as opposed to only unpleasant but inchoate inner feelings. An experiment by Overmeier et al. (1971) supplied some measure of support for this suggestion, since dogs trained to avoid shock by nosing a panel to their left at one signal, but to their right at another, did so more quickly if each signal consistently predicted shock to a particular leg (whether on the same or opposite side as the response needed to avoid it) compared with animals for which either signal was followed by shock to either leg at random. The authors of this report argue that signals predict something specific, and not only something which is quite generally a bad thing (cf. Trapold, 1970, p. 155). There is evidence that rats also, as well as dogs, may acquire reactions to a signal that are specific to particular signalled aversive events:

Henderson et al. (1980) found that prolonged associations between a signal and an airblast meant that the signal had little effect if was turned on while animals were responding to avoid electric shock, whereas a similar brief association, or a prolonged association long past, meant that responding to avoid electric shock was more vigorous. The suggestion is that one aspect of the association, which was the more long-lasting, was arousal or diffuse fear produced by the signal, but a second, more ephemeral part of the association linked to the signal reactions or representations specific to the airblast.

However, whatever the level of cognitive representation of specific feared events that may take part in avoidance learning, it is by definition the instinctive and built-in reaction to both the aversive events themselves and to unpleasant emotional states associated with them to withdraw and shrink from them, and to perform any response which lessens contact with them. It is completely possible also that part of the instinctive and built-in reaction to aversive events include modulation of the classical conditioning process, so that an intermittent association between an arbitrary signal and pain is taken more seriously than an intermittent association between a similar signal and food reward. This is more likely for aversive events which produce emotional reactions of very high intensity; Solomon and Wynne (1954) for instance, proposed that traumatic conditioning of anxiety might be irreversible if feedback from the periphery of the autonomic nervous system caused some kind of brain overload. This is rather vague, but some explanation is needed for the empirical evidence that only one or two associations between a signal and a powerfully aversive event may produce indefinite aversion to the signal (Wolpe, 1958; Garcia et al., 1977b). The cases where only one pairing of a signal and an aversive event occurs make nonsense of the otherwise well-supported theory that only a statistically reliable correlation or contingency
and the outcome-assuming nature of withdrawal.

Part of the asymmetry may lie in the central criteria used for conditioned emotional associations, with, as it were, more stringent internal statistical criteria required for hope than for fear. But Solomon and Wynne (1954) may well be right to point to the autonomic system as well. Aversive stimuli arouse the ‘fight or flight’ or ‘behavioural inhibition’ syndromes of the sympathetic nervous system and limbic brain system respectively (Gray, 1982). Once this sort of physiological arousal has reached a certain point, it may become aversive in its own right, and thus a signal initially associated with a strong aversive stimulus may become motivationally self-sustaining (Eysenck, 1976; Walker, 1984, chapter 9). This is still somewhat speculative. There is little doubt, however, that many of the peculiarities of avoidance learning, and indeed of any form of reaction to aversive stimulation, can be attributed to the instinctive behaviours of the species involved, or ‘species-specific defensive reactions’ or ‘SSDRs’ (Bolles, 1970, 1978). It may be that fearful emotional states generally produce more rigid and reflexive behaviour than relaxed exploration or systematic foraging for nutritional necessities. In any case, it is possible to point to many specific responses, such as leaping and frightened running, or passive crouching (freezing) by laboratory rats, which, are automatically elicited by particular aversive stimuli, and therefore are likely to occur in response to associated signals for such stimuli whether they are useful or not. This has led to assertions that many kinds of learning induced by aversive stimuli are special cases, and not explicable in terms of general principles, but it is not necessary to abandon all general principles provided that included among them are principles which take into account instinctive behaviours. Taste-aversion learning is a case in point.

**Taste-aversion learning**

It is perhaps a measure of the persisting effects of the dust-bowl empiricism of pre-war learning theories that the phenomena of taste-aversion learning should initially have been found surprising, and that the concept of natural functional relationships between stimuli in learned associations should have taken so long to take root (Garcia, 1981). It is a fact of life that though eating is essential, eating the wrong thing can be disastrous, and to the extent that the process of learning is useful in foraging and food selection, animals, especially those with a varied diet, ought to be capable of learning from experience foodstuffs that are best avoided. Thus young jays quickly learn not to eat moths with a foul taste on the basis of visual cues, and it is well known that some species of moth which are palatable have evolved markings like those of others which are not, because of selective predation a phenomenon known as mimicry (Maynard-Smith, 1975). The biological advantage of taste lies entirely in distinguishing what should be eaten, and when, and how eagerly, from what should not be eaten; and these categories, although they may be to some extent innate, might usefully be modified according to the post-ingestional consequences of specific eating experiences. There is a certain amount of evidence that, as Hull (1943) would have predicted, metabolic usefulness of what is eaten may lead to slight alterations in taste or smell preferences for example protein-deprived rats increase their preference...
for the odour of a diet associated with receipt of balanced proteins (Booth and Simpson, 1971; Green and Garcia, 1971; Holman, 1968; see also Booth et al., 1972; Booth et al, 1976). There is much clearer and stronger evidence that animals very rapidly become averse to the taste of a food eaten before they became ill.

This became apparent in studies of the effects of radioactivity on animal behaviour. Exposure to radiation affects the intestinal tract and makes animals ill; after they have recovered they may refuse to eat foods previously consumed (Haley and Snyder, 1964; Rozin and Kalat, 1971). This would not have surprised Pavlov in the least, but a considerable stir was created when Garcia and Koelling (1966) published a careful experiment which suggested that the effect was selective, in that taste cues much more than visual cues appeared to acquire associations with illness. The experimental technique was to place thirsty rats in a small box containing a drinking spout for 20 minutes a day, measurements being taken of the number of times they lapped at the spout. The water might be given a sweet or a salty taste, and an attempt was made to provide audiovisual feedback of a roughly equivalent kind by arranging that a flash of light and a click would occur each time a rat lapped at the drinking spout. Rats were first pre-tested to assess rate of drinking ‘bright-noisy-tasty’ water under these conditions. During the training phase, on most days the animals were allowed to drink plain water undisturbed, but every three days the distinctive sight, sound and taste feedback was given, and the animals subsequently became ill, some because, while drinking saccharin-flavoured water, they were exposed to a sufficiently strong dose of X-rays, and others because lithium chloride, which tastes salty, was added to their water. For comparisons, yet other rats were allowed to drink bright-noisy-salty water while a gradually increasing shock, which eventually suppressed their drinking, was applied to the floor (‘delayed shock’), and a fourth group had alternate 4 minutes of immediate shock when they drank with the three kinds of feedback, but no shock when they drank plain water without audiovisual feedback.

All the animals in this experiment (Garcia and Koelling, 1966) had very much suppressed drinking by the compound cues during the training procedure, the shock animals somewhat more than those poisoned. The crucial phase occurred when no further aversive events, either sickness or shock, were imposed, and the rats were tested separately with water that had the previously experienced taste, but no sight and sound feedback, or the sight and sound feedback with plain water. These tests showed very clearly that animals which had been shocked drank just as much of the flavoured water as they had done of plain, but drank less when plain water had the sight and sound cues. And by contrast, animals poisoned drank normally under these conditions, but drank very little of water flavoured with saccharin or with normal salt, in the absence of the light and click feedback (the sweet and salty taste having been used for the X-ray and lithium chloride groups respectively). Since all the rats received both taste and audio-visual feedback during conditioning, it would appear that there was a selective tendency to connect the internal and visceral sensations of illness with the cue of taste, and to
connect the pain coming from the outer environment with some aspect of the audio-visual compound. Internal consequences were associated with the internal cue of taste, while external effects were associated with the external modalities of sight and sound. This kind of result has been very widely replicated (e.g. Domjan and Wilson, 1972; Miller and Domjan, 1981; Revusky, 1977). However, the explanation which should be given for this fairly straightforward finding has been a matter of dispute (Milgram et al., 1977; Logue, 1979). It is first necessary to emphasize that the phenomenon is quantitative rather than qualitative. Pavlov (1927) reported that symptoms of illness could readily be associated with the sight of the syringe which normally preceded their induction, in dogs (see p. 73), and rats will become averse to black or white compartments (Best et al., 1973) or to a compound of external cues that represents the particular box they drank in before being poisoned (Archer et al., 1979). As noted in Chapter 3, extremely specialized metabolic reactions, such as those which happen to prevent the analgesic effects of morphine, are capable of being conditioned to external cues which characterize a particular room. And on the other hand, experiments such as that of Logan (1969) quoted above, indicate that peripheral electric shocks (rather than only illness) can alter food preferences. Thus there is no need to assume that certain forms of unpleasant experience can be associated only with biologically appropriate cues. The effects are a matter of degree: what has to be explained is a kind of selectivity, in which when there are several possible stimuli which could be taken as cues for a biologically significant event, whichever stimulus is most biologically appropriate or relevant is likely to be dominant.

There is little disagreement as to the form taken by the phenomenon, but a variety of views as to what should be concluded from it. Differing amounts of emphasis are given to the innate and built-in aspect of whatever mechanism is responsible. Garcia and Koelling (1966) refer to ‘a genetically coded hypothesis’ which might account for the observed predisposition, and the phenomenon of taste-aversion learning is usually taken to contradict tabula rasa assertions about animal behaviour (Revusky, 1977; Logue, 1979), but less specific forms of innate determinacy, such as a gathering together of internal and external stimuli (Revusky, 1977), perhaps as a sub-example of a principle favouring spatial continuity in the formation of associations (Mackintosh, I 983), have been defended.

Garcia himself has tended to interpret what is now frequently referred to as the ‘Garcia effect’ in terms of innate mechanisms, and has drawn attention to the fact that the taste system is neuroanatomically related to visceral stimuli in vertebrates, since input from the tongue and the viscera both are collected in the brainstem, and there is in fact a particular structure there, the nucleus solitarius, which receives both taste and gastro-intestinal input fairly directly (see Garcia et al., 1974, 1977a, 1977b). There are thus anatomical grounds for expecting that taste should be especially likely to be affected by visceral experiences — more so even than smell, since the olfactory input is to the forebrain, where it goes to the limbic system. There is behavioural evidence (Garcia et al., 1974) to support Garcia’s theory that the olfactory system is used for appetitive food-seeking it supplies information, along with vision and hearing, about objects at a distance, but is more closely connected than they
are with motivational urges to find things which taste good but reject things which taste bad. The taste system of the tongue, according to Garcia et al. (1977a, p. 212), interacts with the limbic (and olfactory system), but is also affected by visceral receptors which ‘assess the utility of the ingested material for the needs of the internal economy’. The evidence is that rats do not appear to associate a smell with illness which occurs some time afterwards (while they do so associate tastes: Hankins et al., 1973), but rats do associate smell with pain, when a food substance is paired with electric shocks (Hankins et al., 1976; Krane and Wagner (1975) suppressed food intake by delayed shocks but did not assess the relative contributions of taste and smell).

There is thus every indication that the ease with which taste-aversions are formed reflects innately determined mechanisms, perhaps even in the form of visible neuroanatomical circuits. How does this affect the theoretical questions of the symmetry of reward and punishment, and the validity or otherwise of the general principle of learning? Garcia et al. (1977b) put the case that there is a symmetry between the dislike of tastes associated with bodily distress and cravings for tastes associated with relief from distress caused by illness or nutritional deficiencies. The onset of illness tends to be sudden, and recovery gradual, which makes dislikes very much more frequent than likes, but if thiamine-deficient rats are given a thiamine (vitamin B) injection after drinking saccharin-flavoured water, they subsequently showed an increased preference for it (Garcia et al., 1967). There is thus support for a ‘medicine’ effect, which is in the opposite direction to the taste-aversion effect. It is not necessarily equivalent in all other respects, but clearly post-ingestional (and post-ingestional) relief from aversive bodily states, and associations with subsequent pleasurable internal feelings, change preferences for ingested substances, very strikingly so in the case of human addictions to alcohol or other drugs.

The charge has been made that taste-aversion learning is a specialized and circumscribed phenomenon, with little in common with other kinds of aversively motivated change in behaviours, and that principles of learning should be assumed to be specific both to particular categories of events to be associated, and specific to particular species (Rozin and Kalat, 1971; Seligman, 1970). This charge can however be effectively refuted, since it is possible to point to many similarities between taste-aversion and other forms of learning, and indeed many similarities between very different species of animal, provided that some principle of selectivity in the formation of associations is accepted, such as ‘preparedness’ (Seligman, 1970) or ‘relevance’ (Revusky, 1977; Mackintosh, 1983) and provided that innate motivational mechanisms and innately determined instinctive behaviours are included as determinants of both learning and performance. Revusky (1977) persuasively argues that if the errors of extreme behaviourism and empiricism are renounced, and it is accepted that ‘from a naturalistic point of view, all aspects of the learning process are innate’ (1977, p.46), then many if not all the phenomena of learning can be subsumed under the extremely general principle that learning ‘has evolved to process information about causal relationships in environments’ (1977, p.46).
Similarly Mackintosh (1983) suggests that ‘a function view of conditioning’ 
would readily accommodate any result showing that ‘a natural causal 
relationship’ is easily learned, but with the rider that ‘To the extent that the 
causal laws describing the world in which we and other animals live are 
generally true, admitting of no exception, so there should be general laws of 
conditioning’ (1983, pp.221—2). In a sense this is Hume’s theory of the 
perception of cause-and- effect, turned on its head, since Hume’s point was 
that what we believe to be causal relationships in the outer world are merely 
subjective impressions based on pairings of events; whereas Revusky and 
Mackintosh argue that the mechanisms which determine how and when an 
individual forms associations based on the experienced pairing of events have 
themselves only evolved because (more often than not) the operation of these 
mechanisms will ensure that learned behaviour will reflect biological truth. 
With this principle to hand, we need not be alarmed if animals learn to 
associate tastes with illness, since the mechanisms of learning evolved in a 
world in which illness is in fact often caused by ingested food. This still leaves 
us with the job of describing what the mechanisms are, and exactly how they 
operate, but brings in at the start not only biological function, but the 
assumption that some of the details of the processes of learning in any species 
have been tuned to the realities of that species’ natural life.

Taste-aversion learning has therefore been extremely important as a 
theoretical cause célèbre, requiring much more explicit acknowledgment of 
innate determinants of learning than was previously thought proper. But the phenomenon itself is readily 
incorporated into the newly liberated versions of general learning theory, since 
the phenomenon is in fact readily obtainable in a wide variety of animal 
species, and is readily explicable as a special case of the two-process theory of 
avoidance learning. A general account of taste-aversion learning in several 
species, with a common form of explanation, has been provided by Garcia 
himself (Garcia, 1981; Garcia et al., 1977b). An aspect of Garcia’s biological 
approach is a respect for species differences, but these appear to be less 
marked than one might expect. Birds usually have few taste-buds but excellent 
eyes, and one might suppose on these grounds that taste-aversion learning 
should be subordinate to sight-aversion learning in birds. Wilcoxon et al. 
(1971) did indeed find, in a widely quoted study using bob-white quail, that if 
these birds became ill after drinking blue-coloured and sour-tasting water, they 
subsequently avoided blue water more than sour. It is clearly absurd to doubt 
that birds (apart from the aberrant kiwis, which are flightless and nocturnal, 
and use smell) use vision in food selection, but there is evidence nevertheless 
that there is a special connection between taste and digestive upset even in 
highly visual species. Extremely hungry blue jays catch even poisonous- 
looking butterflies in their beak, rejecting only those whose taste has been 
previously followed by nausea. This indicates a certain general primacy of 
taste (Garcia el al., 1977a) and also follows the principle of matching learning 
to biological causality, since butterflies which look dangerous but taste normal 
are safe (Browner, 1969). More surprisingly, large hawks (Buteo jarnaicensis, 
Buteo lagopus) which have visual receptors in their eyes measured in millions, 
but taste receptors on their tongue measure only in tens, also seem to use taste 
as the main cue for aversion to poisonous bait. A hawk used to eating white
mice, and given a black mouse made bitter with quinine, and then made ill with a lithium injection, afterwards seized and tasted a black mouse, without eating it, and only after that refused to approach black mice. However, hawks given black mice which did not have a distinctive flavour, before being poisoned in the same way, required several poisoning episodes (instead of just one) to acquire an aversion, and this took the form of not eating either white or black mice. It thus appears that taste is more readily associated with illness than are some readily distinguishable visual features of food, even for the most visual of vertebrates. Although greater persistence through time of taste cues has been ruled out as an absolutely necessary aspect of taste-aversion learning in rats, which do not vomit when ill (Revusky, 1977), it is very probable that part of the salience of a strong bitter taste for avian poisoning experiences is due either to its prolonged after-effects in the mouth or to its presence during vomiting, which is a reaction seen in blue jays after eating poisonous butterflies, and in hawks after lithium injections.

Taste-aversion learning is thus not species-specific to rats. The result with hawks also contradicts the ecological hypothesis that rats show the phenomenon only because, as omnivores, they are likely to sample a wide variety of possibly dangerous substances (Revusky, 1977). There may be biological dangers in a carnivorous diet, and this would mean that the ecology was wrong rather than the relation between ecology and psychology, but results obtained with captive wolves and coyotes demonstrate that general as well as specific processes may be engaged by taste-aversions (Gustavson et al., 1974). A pair of captive wolves attacked and killed a sheep immediately on the two opportunities they had before an aversion treatment of being given lithium chloride capsules mixed with sheep flesh and wrapped in woolly sheep’s hide. On the next occasion that a sheep was allowed into their enclosure, they at first charged it, but never bit it. Then they became playful, but when the sheep responded with threatening charges, the wolves adopted sub-missive postures and gave way. Similarly, wildborn captive coyotes were deterred from attacking rabbits by being given rabbit carcasses injected with lithium chloride, although for most of them two such poisonings were necessary. By contrast, if laboratory ferrets were repeatedly made ill after they had killed and eaten mice, they did not stop killing mice, even though not only would they not eat the mice that they killed, but their aversion to mice was apparent in retching when mice were bitten, and rejection and avoidance of the dead carcass. Less than one in five laboratory rats kill a mouse put in their cage (but three out of four wild rats kept in a laboratory: Karli, 1956), but those which do will kill very consistently, except if given aversion treatments, when they are rather more flexible than ferrets, since they will stop eating mice if poisoned after eating, but will also stop killing if poisoned after killing without being allowed to eat their victim.

The theory offered to explain taste-aversion phenomena in these various species is a variant of the theory of classical conditioning, discussed in terms of a ‘hedonic shift’ (Garcia et al., 1977a, pp. 300—6; see this volume, chapter 3, pp. 77—80). Both specific metabolic and reflexive reactions (for instance nausea and retching) and more general emotional evaluation on some like-
dislike dimension become shifted to the signalling stimulus, which is usually
taste in the first instance, from the later events of illness. In some species, in
particular the wild canines, attack in a state of hunger is relatively well
integrated with expectations of eating — in the terms of Adams and Dickinson
(1981b), attack is a purposive action rather than an automatic habit. Therefore
in these species an aversion to the goal has a relatively powerful inhibitory
effect on behaviours which lead to the goal. In other species, or at any rate in
domesticated rats and ferrets, the instinctive behaviour of killing is relatively
independent of representations of the taste of the goal, and therefore aversion
to the taste has less effect on responses which happen to provide the
opportunity for that taste. The hedonic shift would be expected to be
associated with the qualitative aspects of the unpleasant experience it resulted
from, but is not always limited in that way, since the wolves with a single
experience of taste-aversion modified their behaviour to the extent of adopting
species-typical postures of social submission at the advance of a now-
unpalatable sheep. The other example often quoted in this context is the
positive social behaviour directed by hungry rats at conspecifics whose
presence has become the signal for food (Timberlake and Grant, 1975). This
result supports the idea that there is a motivational good-bad dimension which
is partly independent of the type of attractive/aversive event experienced,
whether social, oral,

intestinal or tactile. Garcia et al. (1977b, p. 284) include as indicators of the
bad end of this scale ‘conditioned disgust responses’ which include urinating
on, rolling on or burying food associated with illness in coyotes, and a paw-
shaking gesture in a cougar.

Once a motivational shift has taken place, it is conceivable that new motor
responses could be learned instrumentally under its influence — an animal
might learn to press a lever to allow itself to escape from close proximity to
strongly disliked food. There is little evidence to show arbitrary responses
being learned in this way, but, as is the case for many laboratory forms of
avoidance learning (Bolles, 1978), once a conditioned motivational state has
been established, certain instinctive but sometimes goal-directed patterns of
behaviour are likely to be elicited. A superficial similarity between aversive
motivational states established by electric shock and those which result from
poisoning is that both appear to elicit species-specific responses of burying
unwanted objects, although only a limited amount of information is available
on this. Pinel and Treit (1978, 1979) have however confirmed that rats having
received only one strong electric shock from a wire-wrapped prod mounted on
the wall of their test chamber thereafter appeared motivated to cover up this
object, either by pushing and throwing sand or bedding material over it with
the forepaws, when this was possible, or by picking up wooden blocks with
their teeth and placing them in a pile in front of the prod, if only wooden
blocks were available to them. Rats will also bury certain objects (a mousetrap
or a flashbulb) when first exposed to them in a familiar territory, but not others
(the wire-wrapped prod or a length of plastic tubing: Terlecki et al., 1979). Yet
another set of species-specific behaviours which may be changed when
underlying motivational shifts are induced by the artificial means of electrical
shocks is seen in the social behaviour of chickens. Dominance relationships or
‘pecking orders’ in groups of these birds are usually stable over time.
However, when Smith and Hale (1959) rigged contests between successive pairs of birds in four-member groups by staging a confrontation between hungry birds over a plate of food, and delivering shocks to the initially dominant bird whenever it ate or had social interactions with its partner, they found that they could completely reverse the rankings initially observed, and that the reversals lasted for at least nine weeks without further shocks. It is thus arguable that taste-aversion learning, and related alterations in the motivational value of natural stimuli by pairings with other events, rather than weakening theories of learning, add to their generality by demonstrating that natural and instinctive behaviours are subject to learned change, as well as arbitrary or more flexible responses such as pressing a lever, or running through a maze.

**Stress, learned helplessness and self-punishment**

I have already had cause to comment on the fact that exposure to aversive stimuli has physiological effects, such as changes in heartbeat and in skin conductivity, which can be used as indices of emotional response, and which may thus be useful in assessing the degree to which emotional reactions to aversiveness have become conditioned to prior stimuli. There are a great many other kinds of physiological reaction, induced by exposure to aversive stimuli, including release of adrenaline and corticosteroids by the adrenal glands, and also changes in brain biochemistry, for example the release of natural opioids (Maier *et al.*, 1983; Seligman and Weiss, 1980). Many of these reactions, which are part of the body’s defence against damage and disease, can usefully be subsumed under the term stress (Selye, 1950). Physiological stress is an example of an asymmetry between the motivational systems of reward and punishment. It is possible to consider emotional excitement representing hope, elation or satisfaction as being physiological arousal similar to fear, though opposite in its affective value, and conceivably corresponding positive reactions for extreme fear can be found in sufficiently intense cravings for food, drink, drugs and socially and sexually attractive goals. But there is no departure from physiological normality due to the experiencing of attractive events which serves as a counterpart to the changes under the heading of stress which can be produced by exposure to aversive stimuli.

Stomach ulceration and loss of weight in rats and other mammals is a relatively indirect way of measuring stress, but serves to indicate long-term effects. Measurements of ulceration, together with behavioural evidence, suggest that there are psychological factors in the stress produced by externally painful experience, even in rats. Seligman (1968) found that unpredictable shocks, randomly interspersed with visual or auditory stimuli, produced extensive ulceration in rats, as well as profound suppression of food-rewarded lever-pressing. Control groups receiving exactly the same physical intensity of shock, and the same audio-visual stimuli, but with the shock signalled by these cues, formed no ulcers, and kept up their usual levels of food-rewarded behaviour, except in the presence of the shock signals. Ulceration is also less in rats which receive shock only in the absence of their own avoiding response, than in animals receiving identical physical stimulation which is uncorrelated with their own behaviour, and not otherwise
predictable (Weiss, J.M., 1971). Not surprisingly, in view of these findings, it has frequently been observed that rats will respond so as to be exposed to signalled and predictable, rather than to unsignalled and unpredictable shocks, when given the choice (Lockard, 1963; Miller et al; 1983; Badia et al., 1979).

In monkeys, severe ulceration has been observed even when hardly any shocks are received, if this is only the case because the animals are responding continually for long periods (on a Sidman avoidance schedule, see p. 226) in order to prevent shocks, and may therefore be assumed to be then in a state of constant anxiety (Brady et al., 1958). The existence of this sort of stress response shows both that there may be distinctive physiological changes produced by aversive laboratory procedures, and that fairly complex psychological reactions also occur, particularly involving the predictability of aversive events, and therefore, of course, the predictability of their absence. Degree of predictability of events, especially their predictability on the basis of the subject’s own behaviour, appears to be something which can itself be learned, with a consequent influence on more ordinary forms of learning in the future. This is a conclusion drawn primarily from research into the phenomenon known as ‘learned helplessness’, which has been extremely extensive, due partly to the belief held by some that this kind of learning is an important aspect of

human depression, the most common form of mental illness (see Maier and Seligman, 1976; Seligman, 1975; and Seligman and Weiss, 1980, for reviews). The initial experiments were performed on dogs, using a shuttle-box avoidance test like that of Solomon et al. (1953; see p. 222). Normally in this apparatus, dogs learn within 2 and 3 trials to jump over the barrier as soon as shock is turned on, and eventually learn to jump to the signal before the shock. However if, before this test, dogs are placed in a harness and given at random 50 or more shocks which they cannot escape from, most of them never make even the first escape response, and few if any ever learn to avoid or even escape from the shocks consistently. Seligman’s (1975) argument is that the dogs given inescapable shocks had learned to give up trying, or had learned that they were helpless to escape shocks. Something of this kind may indeed occur, but it is likely that this is not the only consequence of the large number of shocks given in the preliminary treatment. Either a general emotional exhaustion or specific and temporary biochemical changes which inhibit temporarily active learning have been proposed, with some reason, as alternative explanations (Weiss and Glazer, 1975; Seligman and Weiss, 1980).

A third possibility, for which there is strong evidence where rats are concerned (Glazer and Weiss, 1976), though not with dogs (Maier, 1970), is that during the supposedly helpless phase animals are in fact learning passive motor strategies which interfere with later tasks which require highly active behaviour.

The three alternative explanations for the inability to learn which is the phenomenon which characterizes ‘learned helplessness’ are thus: (1) some kind of physiological debilitation; (2) an inappropriate, probably passive, response habit; and (3) a more cognitive set, which in animals must at least amount to a disinclination to appropriately associate response output with desirable consequences, and in people might form part of more elaborate attributional processes, in which helplessness could be connected to beliefs
about one’s own general or specific inadequacies, or about the unyielding cruelties of an unjust and uncaring external world (Abramson et al., 1978; Miller and Norman, 1979; Peterson and Seligman, 1984).

Physiological debilitation

There are certainly temporary after-effects of stressful experiences, which dissipate with time, and which can depress learning. In the first experiments with dogs, Overmier and Seligman (1967) could demonstrate ‘learned helplessness’ in shuttle-box training given within 24 hours of the inescapable shock treatment, but not if there was a recovery period of two days or more. Weiss and Glazer (1975) demonstrated that either shock treatment or exposure to very cold water (2°C) 30 minutes before a shuttle-learning test reduced the performance levels of rats. They attribute this to a temporary depletion of adrenalin-like chemicals in the brain, although since relatively inactive motor tasks were not affected, more peripheral forms of fatigue may also have contributed to the reduction in performance on active tasks. Temporary kinds of exhaustion may thus be important in the early stages of learned helplessness. But they are not the only factor. Dogs which have failed once in the shuttle test, given soon after inescapable shocks, will fail again a month later. On the other hand dogs allowed to learn to escape first, before being given the usual stress of shocks in harness, are unaffected even immediately after the stress (Maier et al., 1969).

Competing response habits

In several experiments (Maier and Testa, 1975; Seligman and Beagley, 1975; Glazer and Weiss, 1976; Jackson et al., 1980) rats exposed to inescapable shocks may subsequently learn a passive response relatively well, but appear to have difficulty in performing a task differing mainly in the degree of activity involved. Therefore it is likely that deficits in readiness to perform very active responses is one of the consequences of inescapable shock treatment. But there must be more cognitive or more associative consequences as well. Maier (1970) showed that even if he explicitly trained dogs to stand still to escape shock, as a preliminary phase, there was little subsequent disruption in their ability to learn the usual active shuttling task. Jackson et al. (1980) observed that pre-stressed rats were just as active as others in running through a Y maze, but nevertheless very slow to learn to turn in the same direction every time to escape from being shocked.

Associative or cognitive changes

Since alternative explanations have limited application, it seems necessary to include a more cognitive explanation of the phenomenon of learned helplessness (Maier et al., 1969). A relatively non-committal way of describing this is to refer to the lack of an expectancy that attempts at active responding will lessen or terminate experiences. More positively, inescapable shocks could result in an animal acquiring the expectancy that shock termination is independent of its behaviour. This interpretation has added weight because of the finding that exposure to a zero correlation between tone
cues and shocks delayed the subsequent learning of an association when the
tone was now paired with shock (‘learned irrelevance’). Mackintosh (1973)
also found that a zero correlation between a tone stimuli and the experience of
drinking water, for thirsty rats, similarly retarded the acquisition of
anticipatory licking when the tone was made a signal for impending delivery
of water. There are thus grounds for believing that an expectational, or
associative, mechanism is affected by the experience of the lack of any
correlation between events. (Maier et al., 1969; Dickinson, 1980). It might be
possible to distinguish the associative aspect of this from the related
motivational deficit or ‘reduced incentive to initiate responding’ (Rosselini et
al, 1982, p. 376). One way of doing this is to show that there are cross-
motivational effects —Goodkin (1976) showed that deficits in the usual task
— shuttling to escape shock — could be produced by exposure to the
relatively unstressful preliminary experience of receiving deliveries of food at
random, irrespective of any organized action by the animals. Inescapable
shocks did not encourage rapid learning of new responses needed to obtain
food in later tests (Rosselini, 1978; Rosselini and DeCola, 1981), and impaired
the subsequent learning by rats of whether they should poke their nose through
a left-hand or right-hand hole to produce food, even when training was long
continued, the correct response being changed (reversal learning). The
deficits in this case lasted for long after the animals had recovered from the
temporary suppression of activity produced by receiving the shocks (Rosselini
et al., 1982). Experience of severe and unrelieveable conditions at an early
counter with aversive stimuli may thus have long-lasting effects on future
behaviour compatible with some kind of reduced confidence in the
effectiveness of action, or in secure regularities of events, but it is worth
noting that the effects of prior shock treatments on the subsequent behaviour
of rats in the experiments quoted above were relatively minor, compared with
the complete disruption of escape-learning in dogs observed by Seligman et
al. (1968) and others.

Self-punishment, discrimination and attention

It needs to be emphasized that while an initial experience of severe and
inexorable painful events leads to later passivity, exactly the same external
trauma has quite different consequences if dogs have already been trained in
their escape task beforehand, in the sense that there do not appear to be any
consequences in this case, since the dogs’ performance on the already learned
task is unaffected, and they go on to escape and avoid normally (Seligman et
al., 1968). Therefore the order of various learning experiences is crucial, and
this is particularly so when strong aversive events are involved —possibly
making another special feature of punishment as opposed to reward. The long-
lasting and counter-productive fixation of initial learning was apparent in the
procedures of Solomon et al. (1953), already described (p. 222). Dogs which
had learned to jump over their hurdle at a signal in order to avoid shocks in a
shuttle-box were undeterred by a new arrangement in which jumping brought
them towards an electrified floor instead of away from it. Some dogs made
anticipatory yelps while jumping, and the experimenters concluded that the
high emotionality caused by the reintroduction of shocks after the dogs had
learned to avoid had strengthened rather than weakened the tendency to jump.
It seems plausible that in instances of this kind, where observational evidence of autonomic arousal was described in terms of ‘symptoms of terror’ (Solomon et al., 1953), the repetition of a previously learned response should be regarded as panic-stricken reliance on first impulses. However, it demonstrates that repeated inescapable shocks (once the animals had jumped on the electrified floor, a gate was lowered behind them) are compatible with highly active responding, as well as with the passivity of learned helplessness.

One argument is that both passivity and jumping away are alternative kinds of natural and instinctive responses to pain, one or other being selected in a very obvious way by variations in procedure, since passive animals have been prevented from moving away from shocks (in many cases by being physically restrained in a harness) and active dogs have been first trained to jump (Bolles, 1970, 1978). This is clearly a major factor, but it is worth also bringing in the difficulty for the animals of distinguishing precisely what might be the best option, especially under conditions of high emotional arousal. Solomon et al.’s animals (1953) had already learned that a signal might be followed by shock. They alternated from one side of the shuttle box to the other, and therefore were shocked on both sides both in early training and in the punishment procedure. Once in a state of fear, they had initially learned that jumping could reduce fear, or was otherwise advantageous in avoiding shocks. Finally, the punishing shocks were of relatively brief duration (3 seconds) and the dogs would have experience of much longer episodes early in training. Therefore it is to some extent understandable that the animals had difficulty in discriminating what was obvious to the experimenters — that jumping, which had once been required in rather similar circumstances, should now be abandoned.

The absence of discrimination between different sources of fear was implicit in the theory of self-punitive or ‘vicious circle’ behaviour originally put forward by Mowrer (1947), derived from the two-process theory of the effects of aversive events. If a response has been learned under the influence of conditioned fear, then punishment, especially if it involves the reinstatement of the original aversive event, may add to conditioned fear, and thus enhance the motivation for the punished response. But particular sources of confusion between necessary and unnecessary activities can sometimes be identified as adding to the likelihood of maladaptive behaviours. Brown (1969) reviewed a number of experiments in which rats ran towards a source of electric shocks, thus exposing themselves to aversive events which they could avoid by not so running. But in most cases it is clear that that activities which are elicited by the aversive stimuli, and also the responses which terminate them, are similar in topography or type to the behaviours which ensure continued exposure. For instance, in the experiment by Melvin and Smith (1967), rats first trained to run down an alley into a safe goal box, in order to avoid receiving shock from the floor of the alley, continued to run (or even started to run again after a period when no shocks were given) when the reality of the apparatus was that the middle section of the alley only was always electrified, and shocks could be avoided more completely by freezing in the start box than by running very fast over the electrified segment of the
runway. The difficulty of distinguishing running towards the safe goal box, after getting shock, and running in the same direction before the shock, presumably contributed to this result.

Attention and aversive events

Since fleeing from danger is ecological necessity for many species in the wild, and even civilized life may be motivated to a substantial extent by apprehension and annoyance, it would be very odd if all forms of learning motivated by even mildly undesirable emotion led to helplessness, depression, or further unnecessary disasters. It should therefore be acknowledged that the phenomena in this section are anomalies, which may reveal an asymmetry between rewarding and punishing motivational mechanisms, in extreme or unusual circumstances, but which but do so only over and above an underlying functional similarity between the ability to seek out pleasure on the one hand and security and safety on the other. In the diagram due to Gray (1975) on p. 215 (Figure 7.2) the symmetry of rewarding and punishing mechanisms is maintained when both add to arousal of some kind. Animals ought to be alerted by receipt of either wanted or unwanted outcomes, even though subsequent learning should be directed at increasing such receipts in one instance and decreasing them in the other. The differing advantages in maintaining alertness to relatively distant negative as opposed to positive outcomes might mean that some species give higher priorities to one rather than the other case. However, the shared advantage of attention to either kind of motivating event may be responsible for the so-called paradoxical effect of mild punishment for correct choices, in increasing rather than decreasing correct choices in certain kinds of food-reinforced discrimination learning (Muenzinger, 1934; Drew, 1938; Fowler and Wischner, 1969). The belief that painful stimuli should motivate learning in general, instead of merely motivating escape, is not altogether without foundation, even though many educational practices based on this belief (for instance the ‘beating of the bounds’ of the City of London, when new apprentices were ceremonially whipped at a series of landmarks as an aid to memory) are very happily discontinued.

Reward and punishment: conclusions

The briefest possible summary is the assertion that rewards are wanted and punishments unwanted experiences, which implies a similarity if not an identity of motivational processes based on attractive and aversive events. However, it would not be surprising if the biological priorities differed as between flight from dangerous or painful stimuli on the one hand, and the pursuit of attractive social or consumable goals on the other. There are in fact both anatomical and behavioural grounds for assuming that strongly aversive stimuli have a greater emotional loading, and a less flexible connection with instinctive patterns of behaviour, than strongly attractive stimuli used in similar ways. But for both attractive and aversive stimuli, behavioural experiments can demonstrate automatic emotional anticipation of significant events, instinctive behaviours released as a result of this, and modification of initial behaviours according to their costs and benefits.

There are thus many similarities between rewards and punishments used as
motivating events in animal learning experiments, and it is arguable that asymmetries between attractive and aversive motivation can be interpreted as matters of degree — unpleasant events merely being more likely to produce conditioned emotional states and associated instinctive reactions than pleasant stimuli of roughly the same motivational weight. This approach would certainly apply to the many experiments in which animals appear to weigh positive and negative outcomes against each other.

However, in the context of severe anxiety and stress, it seems necessary to appeal to special factors which apply to aversive but not to appetitive motivation. Some of these are undoubtedly physiological, and directly related to the reactions of the autonomic nervous system to aversive stimuli. Others may be more cognitive in nature, in the sense that they reflect either instinctive defensive reactions of particular species or more general asymmetries in the processing of attractive and aversive information.