3 Pavlovian conditioning

‘So infinitely complex, so continuously in flux, are the conditions of the world around, that the complex animal system which is itself in living flux, and that system only, has a chance to establish dynamic equilibrium with the environment. Thus we see that the most general function of the hemispheres is that of reacting to signals presented by innumerable stimuli of interchangeable signification’.

The conclusion of Lecture I, Conditioned Reflexes, Pavlov (1927)

Pavlov’s theories

It is a great advantage to be able to discuss the phenomena of habituation with a set of common descriptive terms — to ask how fast the response decrement is, how long the recovery takes, and which alternative stimuli result in dishabituation — even though it is quite inescapable that the mechanisms which cause the phenomena may be as different as the knee jerk and visual pattern recognition. In the same way, it is often useful to describe the phenomena of Pavlovian, or classical, conditioning, as if they were the result of a single set of processes. With the understanding that, since habituation to repeated stimuli may involve very complex cognitive processes, we cannot fix habituation on the bottom rung of a ladder of mechanisms of learning, it is possible to classify Pavlovian conditioning as descriptively slightly more elaborate than habituation, since, at a minimum, it involves two stimuli and one response, instead of only one stimulus and one response. Much follows, however, from this difference between one and two. Adapting to environmental change by ceasing to respond to repeated events may be useful, but it involves a very limited form of behavioural change, no matter how elaborate the encoding of the information received about the events. If, because two different stimuli are associated in the environment, a response to one is given to the other, then, in the first place, the animal involved has increased rather than decreased its behavioural repertoire. And, theoretically, a device has been found which glues together any two sensory experiences, allowing in principle for the assembling of parts into wholes, for the detection of causal relationships, and for the reconstruction of indefinitely long sequences of mental representations.

At any rate, this was how Pavlov himself invariably presented his work on conditioned reflexes (1927, 1955). By the mechanisms studied as conditioning reflexes, ‘groups of various agents or elements of nature, both simultaneous and consecutive, are synthesized by the animal into units. In this way synthesis is effected in general’ (1955, p. 273). ‘Thus, from the point of view of conditioned reflexes the cerebral hemispheres appear as a complex of analyzers, whose purpose is to decompose the complexity of the internal and external worlds into separate elements and moments and then to connect all these with the manifold activity of the organism, (1955, p. 300). The mechanism responsible for the conditioned reflex corresponds to what Helmholtz termed ‘unconscious inference’ (Pavlov, 1955, p. 215). What goes on the brain of the dog is ‘higher nervous activity’ as opposed to the ‘lower nervous activity’ of the spinal cord, and thus:

In the long run, the cerebral hemispheres of the dog constantly affect in the most varying degrees both the analysis and synthesis of stimuli coming to them, and this can and must be termed elementary, concrete thinking. And it follows that this
thinking is responsible for the perfect adaptation of the organism, for its more
delicate equilibration with the environment. (1955, p. 274 original italics)

As these quotations, especially the last one, show, from one point of view Pavlov
had what would nowadays be called a

very cognitive approach to the theory of conditioning. However, he did not neglect to
mention the ecological function of these processes. The point of the conditioning
mechanism was that it supplied ‘a much more detailed and specialized correlation
between the animal and its environment than is afforded by the inborn reflexes alone’
(1927, p. 16). The obvious examples of the evolutionary advantages of this were given
by the ability of arbitrary and distant stimuli to evoke ‘the reflex of seeking food’, and
also ‘the reflex of self-defence’, which arises because ‘The strong carnivorous animal
preys on weaker animals, and these if they waited to defend them-selves until the teeth
of the foe were in their flesh would speedily be exterminated’ (1927, p. 14).

Pavlov’s treatment of conditioning was therefore specifically cerebral rather than
merely of the knee-jerk reflex kind, but his physiological terminology of conditioned
reflexes, when taken up by Watson, Skinner and Hull, among others, was used in a
much more mechanical way. As we shall see, evidence of classical conditioning can be
obtained from a wide variety of animal species, and from several different kinds of
human behaviour.

*Pavlov’s experiments*

The essential features of Pavlov’s experiments on conditioning in dogs are very
well-known. The response measured was salivation, and the main experimental result
was that dogs would salivate to a buzzer or bell which was given as a signal for a few
seconds before food was presented. Given the very general theories which Pavlov put
forward about animals adapting to their environment, as illustrated by the quotations
above, the concentration on experiments where the main focus of interest was the
activity of the salivary glands seems rather surprising, but it is completely explicable in
terms of Pavlov’s original interest, as a physiologist, in the process of digestion (see
Boakes, 1984; Gray, 1979). In the lecture he gave in 1904 when he received the Nobel
prize for his work on digestion, Pavlov said that it was quite unexpected that, in the
course

of his research, he discovered that psychological factors had such powerful effects:

generally speaking, the outstanding role of psychological stimulation in the
processing of food in the digestive canal has not met with proper
acknowledgement. Our investigations forced us to bring these influences to the
fore. Appetite, the craving for food, is a constant and powerful stimulus to the
gastric glands. There is not a dog in which skilful teasing with food does not evoke
a more or less considerable secretion of juice in the empty and hitherto inactive
stomach. (1955, p. 141)
Pavlov’s experiments were very carefully controlled. The dog was isolated from the experimenter, who delivered stimuli by means of automated devices, and recorded quantitative measures of response. After Asratyan (1953).

In order to study psychological effects on digestive secretions more thoroughly, it was enough to measure the effects of selected artificial stimuli on volume of salivation, and this could be accomplished by means of a minor operation to lead salivation out through a tube in the dog’s cheek. The experimental methods adopted in Pavlov’s laboratories were very systematic: the dog was usually separated from the experimenter, and from distracting outside noise, in a sound-proofed compartment, and pneumatic or electrical methods were used to control the delivery of the food and other stimuli (see Figure 3.1).

The acquisition of conditioned reflexes

A wide variety of experimental stimuli were used in Pavlov’s experiments bells, buzzers, pure tones, and the sounds of musical instruments, presentation of illuminated visual patterns, or the sight of gradually rotating objects. None of these artificial stimuli would normally induce a hungry dog to salivate, but if any of them were to be presented consistently for a few seconds before the dog was given food, then the ‘conditioned stimulus’ would by itself elicit copious salivation. But this is not a selective process confined to the salivary gland. As Pavlov put it, ‘the motor component of the food reflex is also very apparent in experiments of this kind’ (1927, p. 22). That is, the dog would turn and move towards the place where it normally got food, and lick its lips. Pavlov referred to the conditioned stimulus as a signal for food, and could be called a proponent of the ‘stimulus-substitution’ theory, since he emphasized that the signal had the same effects as the food — for instance if the beat of a metronome was the signal ‘the animal reacts to the signal in the same way as if it were food; no distinction can be observed between the effects produced on the animal by the sounds of the beating metronome and showing it real food’ (1927, p. 22). Pavlov’s explanation for these new effects of artificial stimuli was in terms of the formation of new physiological paths or new connections within the brain. The analogy he used was the now familiar one of the
telephone switchboard: he had a private line from his home to his laboratory this is like a permanent, 'hard-wired' and readily available inborn connection. On the other hand he could call up the laboratory, or other numbers, through the central exchange, and in this case a special new path has to be provided at the switchboard (1927, p. 25). The formation of a new connection for a conditioned reflex could sometimes be done as quickly, after only a single combination of, say, an electric buzzer with food, but at other times, 10 or more combinations of the signal with food would be required before the artificial signal elicited salivation.

**The inhibition and extinction of conditioned reflexes**

The switchboard metaphor is limited, but is very straightforward. A much less direct explanation was given by Pavlov for cases where a conditioned reflex failed to occur when expected. Instead of being put down as a bad connection, such a case would be attributed to a special process of 'inhibition' — a much more specifically physiological theory. ‘External inhibition’ corresponds closely to distraction — if during an experiment a strange smell wafted into the laboratory, or there was an outside noise; or even if the sun went behind a cloud, the dog might prick up its ears, sniff the air, or gaze in the direction of the disturbance. This excitation of the investigatory reflex would have an inhibitory effect on the conditioned reflex — that is, the dog would not salivate as normal when the conditioned stimulus was presented (1927, p. 44). Very generally, Pavlov attributed the absence of behaviours to inhibitory brain processes, and the presence of any activity to ‘excitatory’ brain processes. The most influential case of this is the hypothesis of the internal inhibition of already formed conditioned reflexes, during experimental extinction.

**Experimental extinction of conditioned reflexes**

If a dog has been conditioned with the sound of a metronome preceding food, and then the metronome is sounded repeatedly without being followed by food, there are various reasons for expecting that salivation to the metronome should cease. Most generally, if conditioning is thought of in terms of the signalling or predictive function of the conditioned stimulus (CS), then disruption of the signalling relationship between the metronome and food ought to be indexed by salivary measurements. In terms of the telephone switchboard analogy, a temporary connection should now become unplugged. But neither of these interpretations was used by Pavlov: instead, the cessation of salivation when a stimulus was no longer followed by food was attributed to a neural process of inhibition, which is assumed to have suppressive effects on response output, without changing the neural connection formed when the conditioned reflex was established. The main experimental result which supports this explanation is the phenomenon of ‘spontaneous recovery’. In the demonstration quoted by Pavlov, a dog which salivates to a metronome is presented with the metronome for 30—second periods, at 2—minute intervals, seven times, no food being presented during this period, or at any other point in the demonstration. Though the first time it salivated within 3 seconds of the metronome starting, giving 10 drops altogether, by the seventh time it did not begin to salivate for 13 seconds, and then only produced 3 drops. This is a demonstration of ‘experimental extinction’. (1927, p. 49). However, when, 23 minutes later, the metronome is turned on again, the dog salivates within 5 seconds and gets a score of 6 drops in the 30 seconds. This is a significant recovery, and because ‘extinguished reflexes spontaneously regenerate in course of
time’, the initial extinction cannot have been due to ‘disruption of the respective nervous connections’ (1927, pp. 59—60).

Another result quoted by Pavlov is what he called ‘secondary extinction’. It was common in his experiments to establish more than one conditioned stimulus at a time, partly since changes of this kind increased the alertness of the animals. Thus a dog might have dilute acid squirited into its mouth after the sound of a metronome, or after the sound of a buzzer, or after the tactile stimulus of a touch on the skin. Now if the metronome is given without the reinforcement, responses to the other two stimuli are extinguished as well (1927, p. 55). Pavlov found it profitable to discuss such cases of interactions between stimuli, as well as spontaneous changes over time, in terms of a labile and diffuse form in inhibitory brain activity. The concept of inhibition as a kind of brain function has proved to be long- lived (see Gray, 1979, 1982), and phrases such as ‘inhibitory conditioning’ are still in frequent use (e.g. Mackintosh, 1983).

**Conditioned inhibition**

A further kind of interaction between different conditioned stimuli is also consistent with the concept of a generally suppressive process. Suppose a metronome again normally signals food, but whenever a whistle is sounded along with the metronome, food is withheld. Accurate expectations would be aroused if the metronome alone was taken to be a positive signal, and the combination of whistle and metronome simply ignored. However, experiments in Pavlov’s laboratory suggested that the whistle was not merely neutral, since its effects could immediately transfer to other combinations. In the case discussed (1927, p. 77), a dog given food after the metronome was also given acid (which also elicits salivation) after a tactile stimulus. When, after training when the addition of the whistle to the metronome signalled absence of food, the whistle was added to the tactile stimulus for the first time, salivation was almost completely suppressed. In all the experiments of this kind, it had to be shown first, of course, that the sounding of stimuli such as the whistle did not suppress salivation by distraction, but many examples confirmed that a stimulus which suppressed responding when combined with one positive signal tended to suppress salivation to any other signal. Thus Pavlov called additional stimuli of this kind ‘conditioned inhibitors’ (1927, p. 77). The procedure of assessing a stimulus with suspected inhibitory properties by adding it to another with known response-eliciting potential is still in use, and indeed Mackintosh (1983, p. 178) suggests that, under the sobriquet of ‘the summation test’, it is ‘the single most useful measure of inhibitory conditioning’.

**Synthesis and analysis in generalization and discrimination**

Yet further areas of current research where Pavlov established the experimental procedures and some of the technical terms involve the effects of conditioning on perception. A separate chapter (chapter 8) will be devoted to modern research on these topics, but it is worth noting here some of Pavlov’s assumptions which have been neglected, mainly the distinction between analysis and synthesis, and the corollary that perceptual complexity should be one of the dimensions which research can and should uncover. By ‘analyzing mechanisms’, Pavlov meant partly just what we would nowadays call the sensory systems, of vision, touch, hearing and so on, but also the involvement in these of attention, that is something ‘which selects out of the whole complexity of the environment those units which are of significance’ (1927, p. 1 10). Synthesis was the process ‘by means of which individual units can be integrated into an excitatory complex’, something which would more likely
be referred to now as cross-modal perception, object perception, or the formation of schemata, representations, or internal descriptions (Walker, 1983a). Pavlov’s constant reference to analysis and synthesis may merely be a consequence of the fact that as a young man his favourite author was Herbert Spencer, who also used these terms frequently, but the experimental facts of discrimination learning, no less now than then, require some roughly similar theoretical attempt to account for attention and pattern recognition (Sutherland and Mackintosh, 1971; Sutherland, 1960).

A more detailed account of Pavlov’s experiments in these areas will be given in later chapters, with further discussion of this theoretical point, but for present purposes it is important to note that Pavlov carefully distinguished ‘elementary’ analysis and synthesis from ‘higher’ types: the former being associated with the capacities of the sense organs themselves, as with absolute thresholds for pitch or visual acuity, and the latter being what he took to be central perceptual processes — for instance the integration of information of both ears to compute localization of sound sources. It was very evident to Pavlov, though it has often been forgotten in subsequent discussions of Pavlovian conditioning, that ‘only with the progressive development of the analyzing activity of the nervous system is the organism enabled to multiply the complexity of its contacts with the external world and to achieve a more and more varied and exact adaptation to external conditions’ (1927, p. 111). That is, Pavlov believed that the scope and limits of perceptual functions would be different in different animal species, and indeed at different levels of the mammalian nervous system, and that these differences would be made clear by conditioning experiments (1927, p. III).

Lesser systems

In 1931 Clark Hull, President of the American Psychological Association in 1936 and eventually to become one of the most famous stimulus-response theorists of all time, was co-author of two technical papers which described simple electrical circuits which displayed ‘fairly accurately the behaviour of the mammalian conditioned reflex’, it being made explicit that this was done since ‘The conditioned reflex is the basic mechanism of more complex mammalian behavior’ (Krueger and Hull, 1931, p.266; Baernstein and Hull, 1931, p. 99). One of these circuits is shown in Figure 3.2.

Figure 3.2. The conditioned reflex modelled as a simple battery circuit

The large battery, E, on the left means that the switch Su can be regarded as an unconditioned stimulus turning on the lamp, L, as the unconditioned response. Small rechargeable batteries, E1 to E5, allow the switches S1 to S5 to act as conditioned stimuli. These will also turn on the lamp, L, but only if they have previously been turned on at the same time as the main switch, Su. The physical characteristics of this type of circuit can be chosen so that the effectiveness of any of the secondary switches will increase gradually with repeated ‘conditioning’ trials, decrease with ‘extinction’ trials, when they are used on their own; and exhibit ‘generalization’ due to leakage through the resistors R1 to R4 from adjacent batteries. After Krueger and Hull (1931).
Metaphorically, ‘push buttons were the sense organs’, the ‘flashing on of a light the 
response’, while ‘copper wires serve as nerves’. In figure 3.2, the learning elements are 
storage cells or capacitors which are charged from the main battery when the push 
button switches of the unconditioned and conditioned stimuli are depressed 
simultaneously. With repeated conditionings of this kind, there is orderly generalization 
of the conditioned response through the resistors. Experimental extinction occurs if the 
conditioned stimulus switch is used often enough by itself, but, due to the characteristics 
of their storage batteries, Krueger and Hull were able to demonstrate a certain amount of 
’spontaneous recovery’ by giving the circuit a rest.

By comparison with modern electronics, this circuit is elementary in the extreme, 
but it serves all the better to make one of the points as originally intended — that several 
of the phenomena demonstrated by Pavlov with the salivary reflex of dogs require, in 
their most basic form, only the most rudimentary of physical mechanisms for their 
reproduction. In terms of behavioural evolution (see chapter 9), it is surely important to 
remember both that the neural control of some very adaptive behavioural processes need 
not be terribly complicated, and that if we describe ‘classical conditioning’ as a single 
behavioural process, it must be understood that the process includes both some very 
straightforward kinds of association and some very much more complex perceptual 
abilities, such as those demonstrated by Pavlov’s dogs when they became conditioned to 
the sound of a particular musical instrument, a certain arpeggio played on any 
instrument, or the sound of their own name (see chapter 8). We should not expect any 
simple electrical circuit to do these more complicated things.

Neither should we expect very complicated perceptual processes from Aplysia 
californica — but we ought not to be surprised if, within a small range of stimuli and 
responses, it accomplishes the sort of conditioning that can be achieved by a row of 
batteries. This result has recently been reported by Carew et al. (1981, 1983). As is the 
case with most artificial systems, in Aplysia the conditioned stimulus tested already had 
a connection to the conditioned response wired in: the 
conditioned stimulus was a light tactile stimulation of the siphon of the animal (see 
chapter 2, pp. 47—8) which normally produces a weak withdrawal reflex of the siphon 
and gill. The unconditioned stimulus was an electric shock to the tail, of sufficient 
strength to produce a very vigorous withdrawal reflex including the siphon and gill. 
Thus, if the siphon is weakly stimulated just before the strong electric shock is applied, 
15 times, at 5—minute intervals, then 30 minutes later a much longer than usual 
withdrawal response is given to the tactile stimulus to the siphon (delivered with a nylon 
brush) tested by itself. As this effect is less than compelling evidence for learned 
behavioural change, an important comparison is that there is no lengthening of response 
to siphon stimulation if this stimulus is given in the intervals between the strong electric 
shocks to the tail, these being signalled by a weak electric stimulus to the mantle-shelf. 
In this case it is the response to mantle-shelf stimulation which is lengthened. This 
comparison is at least metaphorically similar to the differential conditioning of a dog, to 
salivate to a buzzer but not to a bell. The crucial anatomical dissimilarity is that the two 
conditioning stimuli travel down different neurons in the Aplysia case, while the dog’s 
brain has the task of distinguishing two stimuli which both arrive down the auditory 
nerve. A major reason for interest in the simpler task of the sea-slug is of course that the 
exact neural changes which take place during conditioning can be more readily studied. 
Hawkins et al. (1983) used rat-sized Aplysia preparations in which the nervous system 
was dissected free from the body but left attached to the tail. Direct stimulation of two 
individual sensory neurons for the siphon could then be used as conditioned stimuli,
either paired or unpaired with electric shock to the tail, the conditioned response being assessed as the excitatory post-synaptic potential (EPSP) produced in a particular identified motor neuron for the siphon.

The model of conditioning in Apysia suggested by the results of these experiments is rather different from the circuit presented by Krueger and Hull (1931), but quite similar in principle to the alternative circuit presented by Baernstein and Hull (1931), in which activity in the conditioned stimulus part (at the heater of a ‘mercury-toluene thermoregulator’) was necessary before the unconditioned stimulus could have an effect. (Hawkins et al. (1983) describe their model as ‘activity-dependent amplification of pre-synaptic facilitation. In this, the unconditioned stimulus not only produces the unconditioned response, but activates a ‘facilitator neuron’, which makes subsequent unconditioned responses more likely, even to unpaired conditioned stimuli (this is known as sensitization). However, if there is activity in a sensory neuron at the same time as the facilitator neuron is stimulated by the unconditioned stimulus, then the ability of that sensory neuron to produce a potential in the motor neuron, on the other side of a synapse, is selectively amplified. This model has yet to be confirmed, but its authors suggest that it could be very general, operating in vertebrate animals as well as in other parts of the nervous system of Apysia (Hawkins et al., 1983, p.404). The general thrust of their idea is conveyed in Figure 3.3. The crucial element is the facilitator neuron (F) which has the built-in capacity to greatly amplify the effect of any sensory neuron at the synapse with the motor neuron. There are cells in Apysia denoted ‘L29 cells’ which are possible candidates, since they project very diffusely and are themselves excited by motivationally significant stimuli. In vertebrates there are diffusely projecting systems of neurons which might do a similar job, but this is purely speculative at present.

Eventually, the detailed theories of how Pavlovian conditioning is physiologically accomplished will presumably become much less speculative, but the work on Apysia now stands as a reductio ad absurdum for the basic processes of association in classical conditioning of the same kind as the simplified electrical circuits discussed by Hull. That there are available neural circuits of the kind portrayed in Figure 3.3 is not in doubt. How such basic modules of association are assembled or utilized to produce central representations of complex perceptual and motor events, and mental associations between them, is of course a different kind of question, despite claims to the contrary (e.g. Hawkins and Kandel, 1984).
F = a facilitator neuron, which makes the connection between cs and R stronger

Figure 3.3 The conditioned reflex modelled as a simple neural circuit.
US is a neuron which is activated by the unconditioned stimulus, and is directly connected to the neuron R, which outputs the unconditioned response. CS is a neuron which has a connection to R, but which does not normally fire R. However, the CS becomes able to fire R if the CS output to R is accompanied by the output from the facilitator neuron, F. Thus the connection between CS and R is made stronger when CS and Us are active at the same time. After Hawkins a at. (1983).

Spinal conditioning
If facilitatory circuits exist in the nervous system of the sea-slug, which enable the associative effects of stimulus pairings to be experimentally demonstrated, we have little reason to be sceptical of claims that the vertebrate spinal cord is also sensitive to temporal relationships between similar reflexes (though we have many reasons for being sceptical about whether the associative capacities of the spinal cord parallel in all respects those of the intact animal, contra Beggs et al., 1983, p. 531). Shurrager and Culler in 1940 reported that the leg-flexion reflex in acute spinal dogs (anaesthetized dogs with severed spinal cords) could be transferred from the unconditioned stimulus of a shock to the hind paw to the conditioned stimulus of a weak shock delivered to the tail, by using this as a signal for the stronger shock. It should be noted that (a) perception of the stimuli was not in any sense a demanding task, and (b) the two stimuli were qualitatively similar, being both electric shocks: these two points also apply to other studies using spinal animals. Shurrager and Culler’s experiments of 1940 were subsequently criticized by Kellogg (1947) and his colleagues, who failed to replicate them, on a number of technical grounds, and the results of conditioning experiments on spinal mammals have remained ambivalent (Patterson, 1976). Beggs et al. (1983) have presented rather more orderly data than usual, but with a highly specialized procedure: the conditioned stimulus was electrical stimulation of the sensory nerve for a particular muscle of the leg (the peroneal) and the conditioned response was electrical activity in the corresponding motor nerve, the unconditioned stimulus being strong electrical shock to the ankle skin of the same leg. (The animals were anaesthetized and also paralysed by spinal cord transection). There appeared to be a gradual increase in the magnitude of the conditioned response over initial pairings, but there are several incongruities in the data. For instance, the magnitude of the response continued to increase if the conditioned stimulus was presented by itself at 10—minute intervals, though it decreased in groups where this was done at 1—minute intervals.
Whether or not much weight is attached to such peculiarities in the results, it is obvious that increments in the response of a motor nerve to stimulation of its sensory partner is not what Pavlov had in mind when he spoke of the importance of conditioning lying in the function of ‘reacting to signals presented by innumerable stimuli of interchangeable signification’ (1927, p. 15). Spinal conditioning of this kind can only be a special case, which may be of value for several reasons, but which can tell us little about what Pavlov called the analyzing and synthesizing activities of ‘the crowning achievement in the nervous development of the animal kingdom’ — the cerebral hemispheres (1927, p. 1).

A curiosity of spinal conditioning is that some of the strongest data supporting the notion of associative connections in the spinal cord has been obtained from a human paraplegic patient. Ince et al. (1978) reported an experiment which was combined with an attempt to enable a patient to regain voluntary, though artificial, control of the bladder-emptying reflex. A self-administered strong electrical shock to the abdomen would elicit the emptying of the bladder, and the intention was to condition this reflex to the much less violent stimulus of a mild electrical shock to the thigh. It was first shown that this mild stimulus initially did not have any effect on bladder-emptying, even after experience of the stronger abdominal shock, thus ruling out sensitization. Then, during seven sessions, the conditioned stimulus was given for 3 seconds at a time, with the unconditioned stimulus, which elicited bladder emptying, always occupying the last 2.5 seconds of this interval. In this phase 54 pairings were given, and subsequently the conditioned stimulus by itself resulted in roughly the same amount of bladder emptying (77 ml. per session) as had the original strong shock used as the unconditioned stimulus. It is unlikely that associative processes of this kind, at the spinal level, are of any great importance in normal human behaviour, but the possibility of associations which, as in this instance, are divorced from the higher kinds of cognitive processing is useful to keep in mind for more intermediate phenomena, such as the metabolic responses discussed below (pp. 73-4).

**Conditioning in decorticate mammals**

Between 1901 and 1909, there was an intense controversy in physiological circles in St Petersburg between Pavlov and his associates on one side, and on the other workers in the laboratory of Bechterev, a very eminent neurologist (Babkin, 1949, pp. 89—94; see Boakes, 1984). The Bechterev side believed that there was a localized centre for salivation in the cerebral cortex, without which no conditioned salivary reflexes could be established, and this was contrary to several results obtained in Pavlov’s laboratory. The conflict was resolved on a famous occasion on which Pavlov and his team went over to Bechterev’s laboratories to observe two dogs with cortical lesions which, it was claimed, had thus lost all their salivary reflexes. The main demonstration apparently consisted of waving a glass jar containing sugar lumps in front of the animals, without eliciting salivation. Pavlov’s reaction was to insist, against all protests, on performing a quick experiment of his own. He demanded a bottle of weak hydrochloric acid and a test tube, and poured acid from the bottle into the test tube, and then into the dogs’ mouths, several times, and then

waited for the salivation which this produced to stop. Now, pouring acid into the tube in front of the dogs consistently elicited more salivation, even though the acid itself was not transferred to their mouths.
This should have ended the controversy over the role of the cerebral cortex in conditioning: it is by no means essential for simple conditioned reflexes, but it is certainly necessary for the life of the normal animal, and for the full range of conditioning results. As Pavlov (1927) summarized it in the first chapter of his book, in a decorticate animal ‘the number of stimuli evoking reflex reaction is considerably diminished; those remaining are of an elemental, generalized nature, and act at very short range . . . finely discriminating distance receptors lose their power’ (p. 13). Subsequent research has amply confirmed both these points: conditioned reflexes may certainly be observed in decorticate mammals — but the range of possible stimuli is obviously drastically reduced, and various other abnormalities may appear. For instance, Oakley and Russell (1976) compared normal, decorticated and hemidecorticated rabbits in the conditioning of the eyeblink (nictitating membrane) response (hemidecorticates have the cerebral cortex removed from only one of the two hemispheres). A weak electric shock delivered to the skin near a rabbit’s eye very reliably elicits the protective unconditioned response of closure of the nictitating membrane, or third eyelid. If a diffuse light is turned on half a second before each of a series of shocks, then there is gradual acquisition of the conditioned response of blinking immediately to the light. Oakley and Russell’s procedure was slightly more elaborate, since sometimes a tone was sounded, without being followed by shock. For separate groups of animals, the tone was the signal for shock but the light was not. The results showed that even the totally decorticated animals were just as accurate as normals at blinking to whichever stimulus was the signal, but not to the other, even though they were slightly slow at acquiring the conditioned response. Thus, the differentiation between a visual and an auditory stimulus can be achieved by subcortical mechanisms. Pavlov’s own results suggested that even the differentiation between two notes a semitone apart could be achieved in a dog lacking auditory cortex, although this dog showed no sign of differentiating between an ascending and a descending scale of the same notes, or of recognizing its own name (1927, p. 336).

Taking together these results from spinal or decorticate mammals, and the conditioning phenomena observed in *Aplysia*, other gastropod molluscs, leeches, and various worms, and indeed taking into account the behaviour of simple electrical circuits, it is obvious that the bare bones of reflex association can be accomplished at relatively elementary levels of neural organization (Fantino and Logan, 1979; Sahley *et al.*, 1981).

**Conditioning of metabolic responses**

Continuing with the less cognitive aspects of conditioned associations, we return to various digestive or metabolic activities which are even more internal and less involved with purposive actions than salivation, like the gastric secretions first studied by Pavlov in dogs. Pavlov (1927) passed on the observations of a Dr Krylov of Tashkent, who had repeatedly injected dogs with morphine in the course of medical research. This initially produced a sequence of profuse salivation, and then vomiting, followed by sleep. But after five or six days of this, Krylov noticed that dogs would begin the phase of profuse salivation before he had actually given them the injection, and that they would also continue the sequence of vomiting and sleep without the drug itself. In the most striking cases the sequence began as soon as the dog saw the experimenter, whereas in other animals the effects were only observed if the complete normal procedure took place including wiping the dog’s skin with alcohol, ending with the injection of saline solution (which without conditioning would have no effects). In Pavlov’s own laboratory only
mild symptoms of this kind were seen, but recent experiments on the phenomenon of
taste-aversion learning (see pp. 232—42) confirm that, in coyotes and wolves, external
stimuli which have been preliminaries to drug-induced nausea on only a few previous
occasions will themselves induce vomiting (Garcia et al 1977b).

**Conditioned hypoglycemia**

Another case where external stimuli appear to acquire the properties of injected
drugs is where they precede repeated injections of large doses of insulin. These have the
unconditioned effect of lowering blood sugar levels. The procedure to demonstrate
conditioning (usually with rats as subjects) is to inject an animal several times, in
distinctive circumstances: with the ringing of a bell, (Alvarez-Buylla and Alvarez-
Buylla, 1975) or in the presence of a strong smell, such as mentholatum (Woods, 1976).
Then the animal is injected with saline solution in the same circumstances and blood
sugar level is measured and found to be lowered (see Woods and Kulkosky, 1976, for a
review, and Figure 3.4).

**Neural control of glandular responses and conditioned compensation**

Conditioned nausea, and conditioned lowering of blood sugar, in some ways sound
very straightforward. But the internal regulation of metabolic and hormonal processes is
of course in most cases more involved. In a theoretical analysis of the conditioning of
drug-induced physiological responses, Eikelboom and Stewart (1982) sensibly point out
that what things actually function as conditioned and unconditioned stimuli and
responses is not readily apparent, and suggest that it is inputs and outputs of the central
nervous system parts of complex regulatory feedback systems which must be
responsible for the observed peripheral effects, such as lowering of blood sugar levels.
That the peripheral responses themselves may be symptoms rather than causes of
conditioning effects is indicated by the experiment of Woods (1976), who included a
group of rats which, during conditioning trials, was given a combined injection of
insulin and glucose, with the net result that there was no change in peripheral blood
sugar level during conditioning trials. This did not prevent the group showing a
conditioned lowering of blood sugar when tested with saline solution (see Figure 3.4).

Thus insulin must have some direct action as a stimulus to neural control systems,
which results in neural outputs
Figure 3.4 *Conditioned lowering of blood sugar level.*

Four groups of rats were given the test, at the right of the figure, of an injection of a weak salt solution (saline), and their blood sugar level was measured immediately afterwards. Groups with previous experience of receiving injections of insulin, or insulin with glucose, showed a lowering of blood glucose in response to the saline injection. Control groups with previous injections of saline or glucose did not. After Woods (1976).

... which reduce peripheral blood sugar, and these same outputs apparently become conditioned to non-drug sensory stimuli. That in itself would fit relatively easily into traditional stimulus-substitution theories of classical conditioning (Pavlov, 1927; Woods, 1976; Eikelboom and Stewart, 1982; Stewart et al., 1984). A complication of a more serious order arises because in many cases conditioned effects of external stimuli which precede drug injections are in the *opposite* direction to the unconditioned effect of the drug injection itself. This is true, for instance, of signals which precede low doses of insulin — in this case the signals alone usually increase blood sugar level (hyperglycemia) even though the insulin doses are sufficient to lower it (hypoglycemia: Woods and Kulosky, 1976; Eikelboom and Stewart, 1982). There now appears to be wide acceptance for theories which take the general form of saying that the injection of the drug often produces antagonistic responses to the drug, and that under certain circum-

stances these antagonistic responses, as some of the body’s main unconditioned responses to the drug, become conditioned to external sensory signals. This can be related to Solomon’s ‘opponent process’ theory, which is designed to explain acquired motivational effects (Solomon, 1980; Solomon and Corbit, 1974), but was developed as a specific account of pharmacological tolerance to morphine by Siegel (1975, 1976, 1977; Siegel *et al.*, 1978) and is incorporated into the general account of conditioning of drug-induced metabolic responses given by Eikelboom and Stewart (1982).
Conditioning and morphine tolerance

The experimental evidence in favour of the conditioning factor in morphine tolerance is fairly straightforward, but the inferences from the evidence are slightly less direct than they are in the cases where external stimuli seem to exactly mimic a drug effect. The argument rests on the assumption that, when morphine is injected, there are bodily responses which antagonize the most obvious behavioural effects of the drug, and that it is these unspecified antagonistic responses which become conditioned to external stimuli. This would explain the essential feature of drug tolerance, which is that repeated doses of the same size produce less and less of the main drug effect. And clearly, the first test of this explanation is whether or not external conditions associated with drug administration make any difference to the degree of tolerance observed. Siegel (1975, 1976) assayed the degree of tolerance to morphine by a behavioural test of analgesia. If a rat is placed on a copper plate kept at exactly 54.2°C, it will lick one of its paws, because of the heat-induced discomfort, after about 10 seconds. However, rats given, for the first time, the standard injection of morphine used by Siegel, and placed on the same hotplate, do not lick a paw until they have been on it for approximately 50 seconds, and this is attributable to morphine’s pain-killing effects. If they are given the same dose of morphine, and the same test, repeatedly, then their latency of paw-licking drops down to 10 seconds, indicating that this dose has ceased to have an analgesic effect, that is, to use the conventional term, tolerance has developed to the drug.

Now, of course this test by itself says nothing about how or why tolerance has developed. But if it could be shown that tolerance depends on an association between external circumstances and receipt of the drug, that would be evidence for the conditioning hypothesis. Several experiments, by Siegel and others, supply such evidence. Siegel (1976) gave the paw-lick test in distinctive circumstances in two different rooms. Rats were given a sequence of morphine doses in one of the rooms, and developed tolerance, but then when they were tested in the alternative room, tolerance disappeared (that is, the same dose had a full analgesic effect). Another result which suggests conditioning to external cues is that tolerance does not develop over three doses if these are all given in different environments, although it does if three doses are all given in the same environment (Siegel, 1975). If animals are made tolerant by repeated doses in the same environment, then several injections of saline, instead of morphine, in that environment, reduce the tolerance, which looks like an extinction of conditioning. In fact, the first time rats are given saline instead of morphine in the same room as usual, they appeal to be more sensitive to the heat stimulus than control animals, licking their paws sooner, and Siegel supposes that this is due to the conditioned physiological responses which are antagonistic to morphine producing the opposite effect to morphine, in this test hyperalgesia (Siegel, 1975, 1976).

Conditioning of emotions

It thus seems likely that a complex metabolic and hormonal balance can be influenced by conditioned physiological responses to external stimuli, which is of course consistent with the general assumption that all sorts of bodily conditions are affected by ‘psychosomatic’ factors. There is no reason to suppose that such conditioning effects are limited to drug injections. Cortico-steroid blood levels can be conditioned up to stimuli associated with a poison (Ader, 1976) and down to stimuli associated with daily feeding and drinking (Coover et al.,
1977), in rats, and are known to be influenced by experiential factors in human subjects. Therefore external stimuli associated with natural rather than experimental instances of disease and stress could very well have subsequent conditioned effects. (Russell et al., 1984, report the conditioning of histamine release.) Most gastric secretion and internal biochemical change can take place with little direct psycho-logical effect — we are not necessarily aware of salivation, let alone the detailed functioning of our livers. However, some definite fraction of our body chemistry is intimately connected to the subjective experience of emotion — the adrenal glands supplying the most obvious example, and even here the causal direction or directions of the connection is difficult to unravel (Schacter and Singer, 1962; Maslach 1979) — but clearly it would be surprising if physiological secretions could be conditioned, but subjective emotional changes could not. In most cases, in experiments on human psychophysiology, electrical measurements of bodily functions are used as indicators of more general emotional states. Many such experiments purport to demonstrate direct conditioning effects with human subjects. Increasingly it is being stressed that conditioned effects in people may be modulated by verbal instructions, attribution effects due to these and other variables, and cognitive expectations (Davey, 1983), but it is worth noting some examples where emotional changes observed in experiments with human subjects seem to follow the conditioning paradigm fairly closely. With animals, it is a very powerful theoretical assumption that central emotional or motivational states are influenced by classical conditioning (e.g. Rescorla and Solomon, 1967), but the examination of this assumption will be deferred until chapter 7.

Levey and Martin (1975) report what they describe as classical conditioning of an evaluative response, which is interesting here as an example of a human emotional judgment relatively far removed, one assumes, from glandular secretions. They first gave their volunteers 50 postcard reproductions of paintings and scenic photographs, and asked them to sort through these, rating the postcards as liked or disliked on a scale of —100 to +100. Then, for each subject in the experiment, the two most liked and the two most disliked postcards were pained with randomly selected neutral cards, in a tachistoscope, as if they were conditioned and unconditioned stimuli. Thus the subjects had the experience of looking at neutral cards either just before, or just after, they looked at highly preferred or highly disliked cards, with 20 presentations of each pairing. Subsequently they were asked to sort and rate these neutral cards again, with the finding that the ratings had shifted significantly in the direction of the more affecting scenes which they had been paired with. Thus neutral cards paired with disliked cards suffered a ratings drop of 30 points while those paired with liked cards increased their standing by 16 points. This looks like evaluation by association, which might be an affront to rational aesthetic judgment, but which is certainly not unknown in everyday life.

A rather more direct stimulus for eliciting human emotional change was used in the experiments reported by Ohman and his colleagues on conditioned reactions of another kind to pictures, since the volunteers in these had to submit to the UCS of an electric shock, which is usually confined to the animal laboratory. Shock was delivered to the third and fourth fingers, and skin conductance was measured from the first and second fingers, the unconditioned response being that skin conductance drops when shock is given, but the theoretical interest deriving from the fact that the shock is unpleasant, and may be assumed to induce negative affect beyond the third and fourth fingers. The conditioned stimuli were coloured slides, of either snakes or houses, which were shown to the student subjects for 8 seconds before shock deliveries. Although these produced
no change in skin conductance in control conditions, after only one pairing with the electric shock, anticipatory skin conductance responses were given to the pictures. For pictures of houses, if one to five pairings with shock were given, then presenting the pictures after this by themselves quickly led to the disappearance of the conditioned response (extinction). But with the pictures of snakes (said to be ‘potentially phobic’), the same amount of conditioning was followed by a continuation of conditioned responses when the pictures were given by them- selves (high resistance to extinction: Ohman et al., 1975b).

Varied effects of expectations derived from verbal instructions were illustrated in another experiment, using a similar conditioning procedure, in which for half the subjects, after they had received 10 shocks signalled by a picture of a certain type, the experimenter came into their cubicle, disconnected the shock leads (but not the skin resistance electrodes, which here were on a different hand) and told them that no more shocks were to be given. For the others, at the same point the experimenter merely checked the electrodes. Subjects told that there would be no more shocks showed lower skin conductance responses in extinction, as we might expect, but, with the pictures of snakes, even informed subjects continued to show the conditioned response at an appreciable level, after shocks had been discontinued, and they had experience of this fact as well as the experimenter’s word for it (Ohman et al., 1975a).

How far such artificial conjunctions of events in the laboratory correspond to normal or pathological emotional development in the life span of human individuals is uncertain, but few psychologists now share the view of Pavlov (1955, p. 273) and Watson (1931, p. 158) that direct conditioning is responsible for most or all of human emotional life (Eysenck, 1976; Gray, 1982; Rachman, 1977; Walker, 1984). None the less, it important for theoretical reasons to bear in mind the fact that direct conditioning of human emotions is occasionally possible, even if it is neither frequent nor in every sense fundamental. Experiments such as those described above demonstrate this under controlled conditions — real-life experience of random associations with strong emotions also suggests it, as in illness after certain foods, or when arbitrary geographical or social circumstances elicit distress after associations with rape, burglary or other personal disasters.

**Conditioning and perception**

Most examples of conditioning I have discussed so far have involved events of strong motivational significance, and this is probably not due merely to accidents of experimental convenience — it is reactions to emotionally loaded stimuli which are most likely always to be transferred to preceding signals. However, especially in treating relationships between classical conditioning and habituation, which is not tied to high emotion in the same way, it is necessary to consider whether pairings of complex events, or the constant repetition of stimuli in given sequences, may share in the associative mechanisms uncovered by more conventional experiments. The associationist tradition begun by Hume (see pp. 13—15) and continued by Hartley (1705—57) and J. S. Mill (1806—73; see Boakes, 1984) certainly requires that mental associations should be possible between sensations that do not necessarily impel flinching or salivation. According to this view, active exploration, or even passive experience of sequence, might be recorded as mental links between events or perceptions of them. For instance, associations between sight, sound, touch, smell and taste could be acquired in this way — a kitten learning by gradual experience
that the sound of crumpling paper might be followed by the appearance of a white sphere of a certain size, which will smell not particularly interesting, but which will feel in such and such a way when patted or sensed with the whiskers. This will surely count as a set of associations, and was certainly what Pavlov meant when he talked of inquisitiveness as an investigatory reflex, which assists the animal to achieve equilibrium with the infinite complexities of the world around it; but nevertheless this sort of learning by perception seems a far cry from experiments revolving around reactions to food or shock.

Distinctions between different kinds of motivational system need to be retained, and ultimately it is a matter simply for experimental investigation how much ostensibly different types of learning from experience have in common. But one of the justifications for doing experiments on conditioning, or for developing theories about these, is that the phenomena mean something beyond the details of a single set of data. Therefore experiential effects of unconventional kinds deserve to be considered. It can and has been argued that all our perceptual experience derives from associations between sensations, but experimentally we can only point to oddities. Davis (1976) has reported what he describes as conditioned after-images, obtained by using a sound to signal the brief, bright illumination of a target such as a cross or triangle, observed by human subjects otherwise in total darkness. After about a week of this, at 10 trials a day, subjects report seeing quite vivid after-images if the tone is sounded without the real illumination. Perhaps this ought to be considered as another example when an ‘opponent process’ appears to be more sensitive to conditioning than the first or ‘A’ process (Solomon and ‘Corbit, 1974; Solomon, 1980), since the subjects report the after-image rather than the normally illuminated figure. Another visual after-image, even more like an opponent-process in that it involves colours (and it is the theory off-colour vision that first saw use of the term opponent process, e.g. by Hurvich and Jameson, 1974), is provided by the McCullough effect (McCullough, 1965). For this, the human observer stares at, say, a red vertical grating alternated with a green horizontal grating. Afterwards, if shown black and white gratings of a similar size and orientation, the observer will report that the vertical grating looks green and the horizontal one red. The complementary colour after-effects appear to have become linked to features of the shapes the original colours were experienced with (May and Matteson, 1976). It would seem likely that this is a peculiarity associated with neural adaptation very early on in the visual pathways, but what makes the phenomenon more like conditioning is that it is not necessarily temporary. Holding and Jones (1976) and Jones and Holding (1975) found that if, after the initial experience with coloured gratings, a period of four days intervened before subjects were shown mono-chrome gratings, the effect was still obtained, and indeed a discernible, though weakened, effect was obtainable even after three months. However, the experience of observing the black and white gratings, whenever it occurred, abolished the after-effect. This is superficially similar to the extinction of conditioned reflexes.

Clearly, on any theory of perception, after-effects are special cases, but they serve to draw attention to the extent of association, expectation and context in more ordinary experience, which is often so enormous that it is taken for granted. It is only by illusions and tricks that we are forced to confront the degree to which one sense modality is correlated with another (Day, 1972; Bruce and Green, 1985) in many cases because of habitual associations. If one experiences peculiar sensations of movement when sitting in a stationary train while a train alongside draws out, this is not because of a
genetically programmed correlation between that sort of peripheral vision and that sort of movement, although it would not be surprising if peripheral vision in general was set up to associate self-produced movement with certain sorts of changes. All perceptual effects involved in movement on wheels, and there are many, from riding a bicycle to cornering with the seat of the pants at much higher speeds, are obviously the result of experienced associations between one set of stimuli and another, and cannot be assisted by specialized innate knowledge.

**Classical conditioning in the laboratory rat and pigeon**

Although the conditioning of salivation in dogs has a special place in the history of research on conditioning, this procedure is now very rarely used (see Ellison, 1964; Shapiro and Miller, 1965). With human subjects much data has been gathered with the technique of eyeblink conditioning (Gormezano, 1965), but, for convenience, a large fraction of all research on classical conditioning is performed with laboratory rats or pigeons, inside small chambers known as Skinner boxes. Some of the details of these procedures, and theoretical questions about them, will come up in the next three chapters, but it is necessary to give a brief account of them in the present context. I give the procedural details first, and then discuss the phenomena of backward and second-order conditioning.

**Conditioned suppression of lever pressing in rats**

This provides an indirect measure of conditioning, but is very reliable, and very frequently used. Rats are first trained to press a lever in a Skinner box, for intermittent food rewards (see chapter 5). Then, either while they are doing this, or ('off-baseline') when they are not, because the lever is withdrawn or they are in a separate apparatus, they experience widely spaced electric shocks, each signalled by a conditioned stimulus such as a tone or light, usually of relatively long duration, that is, the tone or light is turned on for 30 or 60 seconds say, and at the end of this a brief shock occurs. The assessment of the effects of this stimulus pairing must always occur while the rat is in the Skinner box, able to press the lever — most simply the stimulus pairings are delivered while the rat presses. What happens is that for unknown but not unexpected reasons, rats reduce the rate at which they press a lever for food reward while the signal for shock is present. The data is usually presented as a 'suppression ratio', that is, the number of lever presses made during the periods when the conditioned stimulus (CS) is present, divided by this number plus however many presses were made during an equivalent period of time without the CS. Thus if the CS has no effect the ratio is 0.5, and if it has maximum effect the ratio drops to 0.0. This means that graphs go down as learning proceeds, instead of up (e.g. see Fig. 4.1, p. 110, Hall and Pearce, 1979).

**Taste-aversion learning in rats**

This could be regarded as the conditioned suppression of eating or drinking. If rats are allowed to eat or drink some recognizable substance (e.g. saccharin-flavoured water) that has been poisoned, or a similar sequence of events occurs because they are made ill by an injection after eating or drinking, then they will consume very much less of that substance when allowed access to it on a future occasion (Garcia and Koelling, 1966). This is true even if they are made ill some hours after the taste experience (Andrews and Braveman, 1975). There has been some argument as to whether this is just another kind of classical conditioning, or a special learning process (it may be both: Garcia, 1981;
Taste-aversion learning has in fact been studied in an impressive variety of animal species (Garcia et al., 1977a; Sahley et al., 1981).

**Autoshaping in pigeons**

As far as the stimulus pairings go, there is nothing to distinguish this from any other kind of classical conditioning — the peculiar name for the procedure arose because the result of it is that pigeons appear to train themselves to peck an illuminated button. The birds are free to move about in the small box, and the illumination of the button, for some seconds (often 8 or 10) is the signal for the presentation of grain, for 2 or 3 seconds, in an illuminated hopper. Hungry pigeons need little training to peck at illuminated grain, and after a few dozen pairings in which the illumination of the button on the wall preceded the presentation of grain several inches below it, they will begin to peck at the button as well. There thus appears to be a transfer of the peck-eliciting attributes of the grain to the stimulus provided by the button. This may or may not be a pure form of classical conditioning (Brown and Jenkins, 1968; Williams and Williams, 1969; Terrace, 1981: see chapter 5) but it is a reliable and useful way of looking at the associative effects of stimulus pairings.

**Autoshaping and food signalling with rats**

If a light bulb is placed inside a translucent lever for rats, or some other method is found of strongly drawing the rat’s attention to something other than food, then autoshaping will occur with rats, that is, if the illumination of the lever signals that food (or water) will shortly become available, the conditioning of positive psychological affect to the lever will result in it being pressed (Boakes, 1977; Wasserman, 1981). This is not as effective a procedure with rats as it is with pigeons, no doubt partly because there is no obviously transferable response pattern (see chapter 5). Using food signalling for rats rather more loosely has, however, proved to be valuable. For instance Holland (1977) presented light or tone stimuli for 10 seconds before the delivery of food pellets to rats while he watched them, and recorded carefully what they did. As it happens the effect of this stimulus pairing was predominantly to cause rats to jerk their heads when they heard a tone which predicted food, but to rear up on their hind legs if they saw the light signal.

**Second-order conditioning**

This has been demonstrated using several different techniques (Rescorla, 1980). If, in using the conditioned suppression procedure, a light signals shock, responding to the light declines (see above). If, after this, rats receive no more shocks, but continue to receive 10—second presentations of the light, now preceded by the sounding of a tone for 30 seconds, then lever pressing in the presence of the tone drops rapidly over the first half-dozen of these pairings (Rizley and Rescorla, 1972).

Holland and Rescorla (1975) used the method of food-signalling with rats, taking a crude measure of general activity. This showed that if 10 seconds of diffuse light preceded the dropping of food pellets, activity in the light increased. Then, in the same experimental boxes, no more food pellets were dropped, but the 10—second sounding of a clicker preceded further presentations of the light, and activity to the clicker was substantially increased as a consequence. In both these experiments, behavioural effects accrued to an auditory stimulus, even though it had never been paired with a motivationally significant event, or rather, even though it had never been paired with a
primary reinforcer, such as food or shock. The auditory stimulus in both cases had been paired with another event, the light, which control groups demonstrated would not otherwise have made any difference but which did make a difference because it itself had previously signalled the primary reinforcer. There would be grounds thus for supposing that the light had acquired secondary motivational significance, and could function as a secondary or conditioned reinforcer (Kelleher and Gollub, 1962) — at any rate, these experiments show that the light was able to serve as a bridge to establish what is known as second-order conditioning, in this case to a tone. Second-order conditioning from a tone to a light is equally possible, and conditioning from one light to another, or one tone to another, works even better, demonstrating that the traditional variable of similarity still facilitates the formation of associations (Rescorla, 1980).

A rather more complicated experiment makes use of the autoshaping technique to demonstrate second-order conditioning, and makes use of second-order conditioning to demonstrate the virtue of consistency in associative learning. Rescorla (1979) first trained pigeons with two alternative stimuli which equally often served as the signal for food. Before food was presented, the response button or key was either lit red on its left half, or lit yellow on its right half. This persuaded the birds to peck the key when it was illuminated in one of these ways, but not otherwise. They were then split into two groups, both of which now received second-order stimulus pairings, in which either horizontal or vertical black lines preceded either the yellow or the red stimulus. As a result, both groups began to peck at the line stimuli. This shows second-order conditioning with the technique of pigeon autoshaping. However, the finding of special interest was that the two groups differed. One of them received consistent pairings in which animals had always vertical then red or horizontal then yellow, or always the other way around, and in the other group all the birds experienced all four possibilities — vertical or horizontal followed by red or yellow. The group which received the more consistent and less variable set of stimulus pairings showed a much stronger second-order conditioning effect, since on the third day they pecked at the second-order line stimuli on 90 per cent of the opportunities, while the birds getting variable and inconsistent stimulus pairings pecked them on only about 50 per cent of their chances.

**Backward conditioning**

One of the rules of first-order conditioning which second-order conditioning appears to share is that backwards arrangements of stimulus pairings are usually ineffective (Rescorla, 1980, p. 9). If a buzzer is sounded just after food has been eaten, this does not lead to the buzzer acquiring the property of eliciting salivation when present on its own in the future, and, generally, behavioural attributes of the first event in a sequential pairing do not transfer themselves to the second very often. However, this is not an inviolable law of all associative learning, as has sometimes been suggested. Common human experience suggests that backwards associations between ideas are at least possible, although Aristotle suggested that it is always wisest, in trying to remember a sequence of things, to take the trouble to find some suitable starting point, and then work one’s way forwards. Experimental studies of human verbal learning (e.g. Horowitz et al., 1964) provide ample quantitative evidence that there are associations backwards in lists of words learned forwards, even though forwards may be better, and rote-learned lists such as the letters of the alphabet or Hamlet’s soliloquies are not easily recited in reverse.
It is possible that backward conditioning is more likely at higher levels of stimulus representation than at lower ones, and that the emotional effects of aversive stimuli transfer more readily to following stimuli than do those of more attractive and appealing events, but no firm conclusions can be drawn at present except that backward conditioning is probably strongly influenced by mechanisms of attention. Using standard techniques, Heth and Rescorla (1973) and Heth (1976) have demonstrated that backward conditioning occurs in the conditioned suppression procedure (see p. 83), in the sense that when 4—second shocks were followed, half a second later, by a 2—second tone-and-light compound, then this stimulus substantially suppressed lever pressing after 10 or 20 pairings, although the effect appeared to dissipate as more and more pairings were given. Rather more dramatic evidence of backward conditioning has been obtained from rats when less conventional techniques of signalling aversive events have been used. Hudson (1939, 1950), as a student of Tolman’s, put rats in a cage in which a striped pattern was mounted on an electrified food cup. After even only one shock from the food cup, his rats, when replaced in this cage a week later, would pile sawdust all over the pattern and withdraw to the other end of the cage, thus demonstrating a conditioned fear reaction (Pinel and Treit, 1978, 1979, and Terlecki et al., 1979, have confirmed this result). What interested Hudson was that the rats appeared to him to look around after the shock, as it were ‘to see what it was that had hit them’ (Tolman, 1948). He then modified the apparatus so that, at the onset of shock, the lights went out briefly while the whole food cup and its distinctive striped pattern dropped out of sight, the rats then being tested for avoidance reactions with the food cup back in place at a later date, as before. Hudson’s own conclusion from the data thus obtained was that:

Learning what object to avoid . . . may occur exclusively during the period after the shock. For if the object from which the shock was actually received is removed at the moment of shock, a significant number of animals fail to learn to avoid it, some selecting other features of the environment for avoidance, and others avoiding nothing. (Tolman, 1948, p. 201)

Hudson also dropped a bundle of pipe cleaners into the cage after the shock had been delivered, finding that the rats then specifically avoided that object, and this aspect of his data was replicated by Keith-Lucas and Gunman (1975). Believing that visually complex 3—dimensional objects were more likely to attract conditioned associations than brief lights or tones, they used as their conditioned stimulus a red rubber toy hedgehog, which flew across the top of the experimental box (on wires) a certain number of seconds after a single shock had been received by several groups of rats, which differed according to this delay interval. Several quantitative measures of behaviour, made subsequently, indicated that if the slow flight of the hedgehog occurred within 10 seconds of the end of the shock, then the presence of the stationary hedgehog in the cage revealed conditioned avoidance.

Gray (1975) quotes a Russian experiment in which dogs got a puff of air in the eye, which elicited blinking, and also had their leg lifted, which meant they would then lift their own leg when it was touched. If these two stimuli were paired together in either order, dogs would afterwards always lift their leg and blink, when either touched or puffed alone, meaning that the response attributable initially to the first stimulus had become transferred to the second. It is conceivable that the relative perceptual vividness and motivational value of the two stimuli in a pairing is one of the variables which influences the greater ease of conditioning to prior, or signalling, stimuli, the lesser of
the two being much more likely to command attention if it appears first rather than second. It is also possible that the factor of relevance to the more powerful event enters into this relationship, via varied attributional or weighting processes. Thus animals may be less likely to associate a subsequent signal with the delivery of food than with distressing experiences, because of the ecological importance of learning to avoid predators ‘seen only after an abortive attack’ (Mackintosh, 1983, p. 210).

**Mental association and conditioning**

For almost every aspect of modern life, from reading to voting, we can be satisfied that experience, and association in one form or another, are important causes of behaviour, by following more or less the same line of argument as that given by Locke. In many cases we might want to talk about purposive action and motor skills rather than pairings of perceived stimuli (see chapter 5), and for most things it would be appropriate to discuss human learning in terms of language and social structures rather than conditioning. But in considering the claims of learning theorists such as Pavlov and Watson, who argued that the conditioned reflex serves as some kind of model or module for more complex forms of experience, we need to say how far the rules which govern learning experiments can be applied to the rest of psychology. It is perhaps safe to say that there is more to human psychology than the rules of conditioning, but it is not clear that the principles of associative learning are therefore irrelevant to human psychology, if the question is put in terms of whether mental associations have anything at all to do with human cognition. Formal mnemonic systems (Yates, 1966; Eysenck, 1977) often use paired associations between numbers and letters or between ideas to be learned and an already known set of vivid images. Learning by rote is still technically possible, though not now in favour as an educational tool, and there are few better examples of the irrational workings of similarity and contiguity as principles of association than those which take place in the process of recollection of material with which to answer the questions set in a traditional 3—hour examination. However, the main connection that it is possible to make between human mental associations and work on Pavlov conditioning is a purely theoretical one. As will be apparent in the next section and in subsequent chapters, the theory of classical conditioning as derived from experiments with laboratory mammals is now firmly mentalistic, in that the locus of associations is between central representations (Rescorla, 1978: Gray, 1979) which may indeed be formed into ‘declarative representation’ that one thing follows another (Dickinson, 1980) in the course of the animal’s attempts to detect ‘a true causal relation between the events to be associated’ (Mackintosh, 1983, p. 222). In this sense it is now believed that, in the case of Pavlov’s dogs at any rate, what was studied as a conditioned reflex in fact reflected a mental association: the dog salivates not as an isolated response, but because it has learned the relation between the events of the signalling buzzer and the subsequent food, and thus salivates when it hears the buzzer because the central representation of the buzzer elicits a mental expectancy of the food. Therefore, in so far as these theories are true and accurate, there is less difference than there was in the days of behaviourist stimulus–response postulates between the laboratory phenomena of conditioning and the laws of association which may apply to human thought.