Chapter 1. Introduction

OVERVIEW

This thesis investigates the role played by regions of the prefrontal cortex and ventral striatum in the control of rats' behaviour by Pavlovian conditioned stimuli, and in their capacity to choose delayed reinforcement.

In this introduction, I will first consider the ways in which animals, particularly rats, represent in their brains relationships or associations between environmental stimuli and motor acts. I will briefly review the evidence that Pavlovian conditioning is a distinct form of learning, and consider the types of associative representations that might be formed during Pavlovian conditioning. Similarly, I will attempt to provide a brief overview of instrumental conditioning, reviewing what is known about the associative representations that influence behavioural responses in the rat, the contribution of Pavlovian learning to instrumental performance, and the complex phenomenon of conditioned reinforcement.

Nearly a century's experiments on normal animals have given us insight into the kinds of associations that form in the *minds* of rats, and it is to be hoped that an accurate knowledge of these associations will assist in trying to understand the way in which rats' *brains* represent and control their world. Although some processes known to psychology may arise from the concerted action of many parts of the brain (e.g. consciousness; Baars, 1988), and some may occur ubiquitously across the nervous system (e.g. memory; Fuster, 1995), neuroscience has achieved considerable successes in mapping some psychological functions (e.g. specific sensory perception, initiation of movement) to distinct anatomical regions or chemical systems in the brain, sometimes fractionating psychological concepts in the process. With this in mind, I will outline the progress that has been made in mapping some of the neural systems responsible for the representations formed during Pavlovian and instrumental conditioning, focusing on the anterior cingulate cortex and the limbic corticostriatal circuitry of which it is a part, and using the perspective of the learning theories described.

Finally, I will consider another complex behavioural phenomenon: delayed reinforcement. Despite its functional importance, the mechanism by which delayed reinforcers control behaviour is not well understood psychologically or neurally, though tantalizing clues have been discovered. I will discuss the relationship between conditioned and delayed reinforcement, and the progress that has been made towards understanding their neural basis.

PSYCHOLOGICAL MECHANISMS FOR ACTION

Unlearned behaviour

Most basic among the mechanisms by which vertebrates influence the world, simple spinal and brainstem reflexes are critical for survival. Such reflexes influence skeletal musculature (respiratory movements, postural reflexes, pain-withdrawal reflexes, and the like) and autonomic function (such as the regulation of heart rate and arteriolar smooth muscle tone to maintain arterial blood pressure). Swallowing is a more complicated example of unlearned behaviour, involving the activation of at least ten different muscles in

a precisely-defined temporal order (Doty & Bosma, 1956). Indeed, innate behavioural patterns can be very complex: an oft-cited example is that of the female greylag goose, which exhibits an innate, species-specific and highly stereotyped behaviour (a 'fixed action pattern') in which it rolls eggs (or any vaguely similar object) into its nest; it will continue the movement even if the egg is lost or removed by an experimenter (Lorenz, 1939; Tinbergen, 1948).

Pavlovian conditioning

Pavlovian conditioning, or classical conditioning, is a term that refers to a set of experimental procedures, in which an experimenter arranges a contingency between stimuli in the world by presenting those stimuli independent of an animal's behaviour. The term makes no assumptions about what is learned.[1] When an initially neutral stimulus (such as a bell) is paired with a unconditioned stimulus (US) (such as food) that elicits a reflexive or unconditioned response (UR), in this case salivation, then the bell becomes a conditioned stimulus (CS) that is now capable of evoking salivation as a conditioned response (CR).

Pavlov, the discoverer of this phenomenon (Pavlov, 1927), argued that a conditioned reflex developed because an association had formed between a representation of the CS and one of the US — stimulus substitution theory (Pavlov, 1927; Tolman, 1934). This would allow novel stimuli, through associative pairing, to control relevant species-specific response mechanisms, extending the usefulness of these responses. However, subsequent behavioural evidence has required the development of theories of conditioning that assume several associative representations are formed during the conditioning process. This evidence is summarized next.

Representations formed during Pavlovian conditioning

Associations are generally believed to be represented in the brain by altering the 'weights' of unidirectional synapses. As synaptic weights can only change on the basis of information available to the neurons involved (the 'locality constraint'), the association of representations A and B in Figure 1 can only occur at points where information about these two representations converge, no matter what mechanisms exist to supervise and use the association. Such associations may be used for different purposes — for example, as a representation of a higher-order property of stimuli (a 'feature detector'), or for commanding behavioural responses directly.

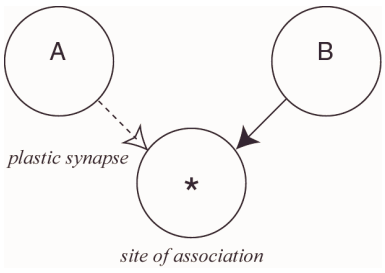

Figure 1. Simple associative representations. Plastic synapses change their weight on the basis of locally available information, requiring convergence of information about the representations to be associated.

In Pavlovian conditioning, there is the potential for several associations to form, as illustrated in Figure 2. In experiments in which the brain is damaged, lesions that removed a representation of either the CS, the US or the response would have obvious consequences not only for conditioned, but also unconditioned, responding (sites A, B, D in the figure). Lesions of site C, representing a central motivational state (such as fear) might not impair primitive unconditioned responses, yet could impair conditioned responses that

[1] Strictly, 'classical conditioning' is the generic term and refers to the experimental procedure described, while 'Pavlovian conditioning' implies Pavlov's interpretation of the process (Mackintosh, 1974, pp. 96/98). This distinction will not be followed in the present discussion, and the term 'Pavlovian conditioning' will be used without implying an underlying mechanism.

were based on the elicitation of fear. Again, however, any properties of the unconditioned response to the US that depended on this hypothesized 'fear' state would be lost.

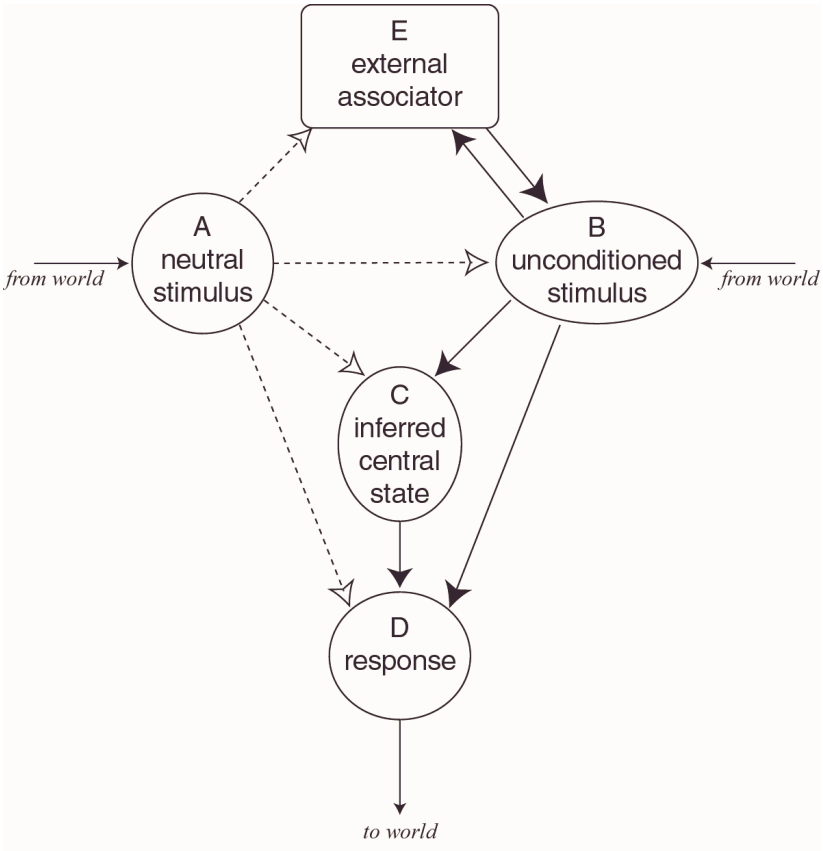

Figure 2. From a theoretical perspective, Pavlovian conditioning has the potential to create associations between a conditioned stimulus (CS) and representations of the unconditioned stimulus (US), central states such as fear, and unconditioned responses. Only a single response is shown; distinctions between different kinds of response are discussed in the text. Dotted lines represent associative links. Bidirectional communication also allows representations to be associated in 'third-party' sites (E) (for similar ideas, see Damasio & Damasio, 1993; Fuster, 1995, p. 88). Note that lesions of such a site might prevent conditioning without impairing any form of unconditioned response, as would selectively disconnecting the CS from a representation involved in responding.

Experimental analysis of Pavlovian conditioning has shown that CS–US pairings may cause the CS to elicit at least three of these representations in the brain (Dickinson, 1980; Mackintosh, 1983; Gewirtz & Davis, 1998). The first and simplest of these is that the CS may become directly associated with the *unconditioned response* (UR), a simple stimulus–response association.

The second is a representation of *affect* — such as fear or the expectation of reward — as demonstrated by the phenomenon of transreinforcer blocking, in which the presence of a CS previously paired with shock can block or prevent conditioning to a CS paired with the absence of otherwise expected food reward (Dickinson & Dearing, 1979). These two reinforcers share no common properties other than their aversiveness and therefore the blocking effect (see Kamin, 1968; 1969) must depend upon an association between the CS and affect. Affective states can therefore be independent of the specific reinforcer and response. This concept has been widely used in theories of learning (Konorski, 1948; Konorski, 1967; Dickinson & Dearing, 1979) and is illustrated in Figure 3. Associations between the stimulus and an affective state appear to be critical in second-order conditioning (in which $S_1 \rightarrow$US pairings are followed by $S_2 \rightarrow S_1$ pairings); unlike a first-order conditioned response (CR), a second-order CR is relatively insensitive to post-training changes in the value of the US (implying that the second-order CR does not depend on $S_2 \rightarrow$US associations) and the response to S_2 may differ from the response to S_1 or the US (implying that it does not depend on $S_2 \rightarrow$R associations) (reviewed by Gewirtz & Davis, 1998).

The third form of representation is *specific to the US*. If a CS is paired with a desirable food and the food is subsequently devalued by pairing it with LiCl injection to induce nausea, so that the food becomes aversive and is rejected, the reaction to the first-order CS changes in normal animals (Mackintosh, 1983).

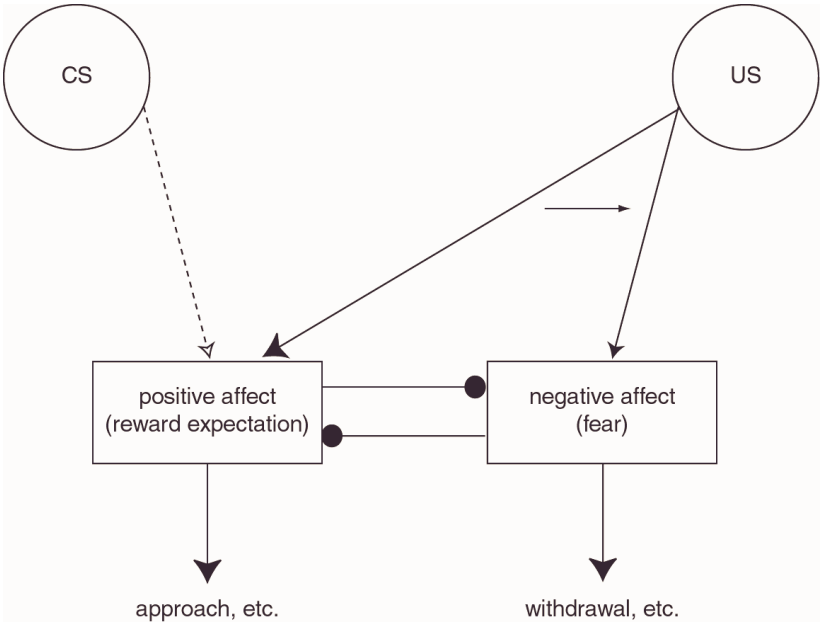

Figure 3. Conditioning to affective states leaves the response independent of the current value of the US. (The CS–affect association corresponds to link A→C in Figure 2.) The CS associates with the affective state elicited by the US during conditioning, but if the US subsequently alters its value (shown as a shift in the US–affect link), the conditioned response (CR) will not alter. The links between the affective states are inhibitory, reflecting the supposition that appetitive and aversive affective or motivational states are mutually antagonistic (Konorski, 1967; Dickinson & Dearing, 1979) (see Mackintosh, 1983, pp. 114–123), though this issue is not important for the present discussion.

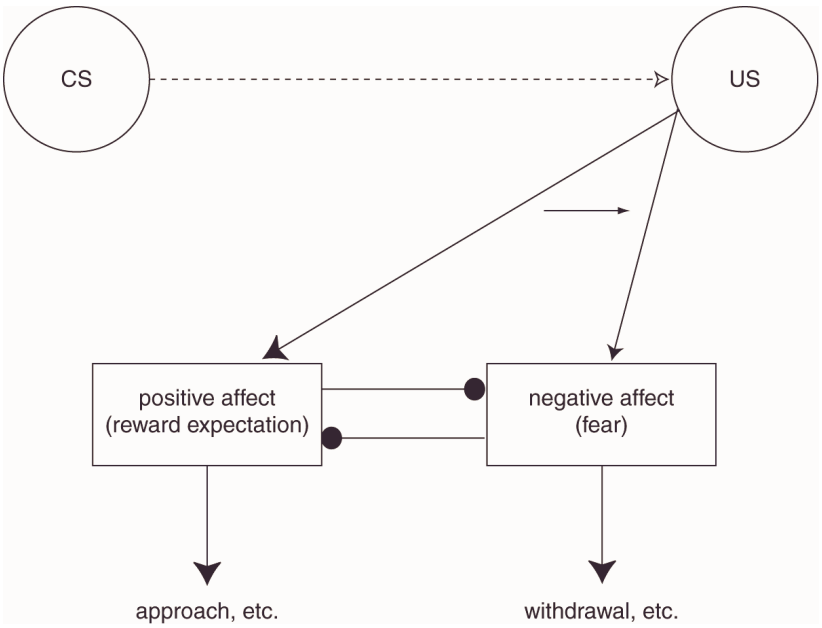

Figure 4. CS–US conditioning allows the conditioned response to alter and reflect changes in the unconditioned response induced by US devaluation. (The CS–US association corresponds to link A→B in Figure 2.) Compare Figure 3.

Therefore the CS cannot have been associated with an abstract 'positive affect' representation, but must have been associated with that particular reinforcer. The association must be specific to the US, because the reaction to a second CS that predicted a different food does not alter, and its connections with valence information must be modifiable and *downstream* of the CS–US association (Figure 4).

Further evidence that a CS becomes associated with a relatively specific representation of the properties of a reinforcer is provided by studies of cued instrumental discrimination (Trapold, 1970). Rats acquired an instrumental discrimination between two levers paired with different appetitive reinforcers more rapidly if the discriminative cues had been paired with the same reinforcers (same condition) in a previous Pavlovian stage than when the outcome was switched between stages (different condition). A rigorous demonstration of the formation of associations between the *specific sensory properties* of stimuli comes from sensory preconditioning (Brogden, 1939), the process by which neutral stimuli are paired in the form $S_2 \rightarrow S_1$, after which $S_1 \rightarrow$US conditioning causes a CR to occur to S_2. Thus, in the first stage, associations form between representations that have no motivational component. Taken together, these pro-

cedures demonstrate that animals are able to encode the relationship between a CS and specific sensory properties of a US and furthermore that they can relate this sensory representation to the affective valence of the reinforcer.

It is likely that further representations are formed during Pavlovian conditioning; for example, it has been argued that animals remember the precise intervals between CS and US presentation, and even that this process — rather than simple associative learning — determines conditioned responding (Gallistel, 1994), though these issues will not be discussed further.

Goal-directed behaviour

The term 'instrumental conditioning' refers to an experimental procedure in which the experimenter arranges a contingency between an animal's behaviour and a reinforcing outcome (Thorndike, 1911). No assumptions are made about the nature of learning; what an animal does in fact learn has been a matter of debate for decades. Although arguments were once made that instrumental conditioning was explicable in terms of Pavlovian conditioning, and vice versa, the two have been doubly dissociated. Classical conditioning is not explicable purely in terms of instrumental response–reinforcer contingencies (for example, Sheffield, 1965, demonstrated conditioned responding even though responding led to the omission of reinforcement). Similarly, instrumental conditioning cannot be explained entirely in terms of classical conditioning; one demonstration was that of Grindley (1932), who employed a form of *bidirectional control* assay. He trained guinea pigs to turn their heads to one side in order to earn a piece of carrot delivered immediately in front of them, and then altered the instrumental contingency, training the same subjects to turn the other way. As there is no differential Pavlovian contingency between the two responses, the asymmetry in responding must have been due to the instrumental contingency. What, though, is learned as a result of this contingency?

Early theorists took the position that the delivery of reward strengthened an associative connection between environmental stimuli and a particular response (Thorndike, 1911; Grindley, 1932; Guthrie, 1935; Hull, 1943). Such learning would represent *procedural* knowledge (Dickinson, 1980), as the structure of the representation directly reflects the use to which the knowledge will be put in controlling the animal's behaviour, and would be inflexible, in that subsequent changes in the value of the reward would be unable to affect responding.

However, it has been shown that rats form more sophisticated and flexible representations in instrumental conditioning tasks. Behaviour may be said to be *goal-directed* if it depends on the twin representations of (1) the instrumental contingency between an action and a particular outcome, and (2) a representation of the outcome as a goal (Tolman, 1932; Dickinson & Balleine, 1994). Simply put, a goal-directed organism presses a lever for food because it knows that lever-pressing produces food *and* that it wants the food. As performance of such behaviour requires these two representations to interact, the knowledge upon which performance is based must be *declarative* — that is, the knowledge is to some degree independent of the use to which it is put. (The interaction is dynamic, and if two actions are to be compared simultaneously, a binding problem occurs — if action 1 is associated with outcome 1, which is associated with high incentive value, that high value must be 'mapped back' to action 1, not action 2 or 3, to encourage its performance. The problem of representing such symbolic processing in a connectionist network like the brain is discussed by Holyoak & Spellman, 1993; Shastri & Ajjanagadde, 1993; Sougné, 1998.)

Instrumental contingency

Free-operant lever-pressing in rats satisfies the instrumental criterion, as shown by a bidirectional control assay (Bolles *et al*., 1980). Not all behaviour may be conditioned instrumentally, however; for example, scratching is difficult to condition in rodents (Shettleworth, 1975; Morgan & Nicholas, 1979). Similarly, the instrumental status of spatially directed locomotion is in doubt, as illustrated by Hershberger (1986) using a 'looking-glass runway' in which chicks were required to run away from food in order to obtain it; the chicks were unable to do so.

Incentive value

The status of an outcome as a goal may be determined by an outcome devaluation test, so long as the test controls for Pavlovian conditioning and is conducted in extinction, so there is no chance for subjects to alter their behaviour by learning a new action–outcome relationship (Dickinson & Balleine, 1994). Using such a procedure, Adams & Dickinson (1981) demonstrated the goal status of food in lever-pressing tasks using hungry rats. The rats were given access to a lever, and one type of food (termed the positive outcome) was delivered contingent upon lever-pressing. Another type of food (termed the negative outcome) was delivered if the rats refrained from responding for a short while. Adams & Dickinson then devalued the positive outcome for one group and the negative outcome for another by pairing it with LiCl injection, and tested the rats in extinction (so there was no opportunity to learn a new response–outcome relationship). It was found that the rats for which the positive outcome was devalued pressed the lever less than those rats for which the negative outcome was devalued. This result cannot be attributed to a general suppressive effect of the devaluation procedure on lever-pressing. As Pavlovian conditioning was also controlled for, the differential effect of devaluing the 'positive' and 'negative' outcomes must have been mediated by the instrumental contingency, a result confirmed by Colwill & Rescorla (1985) using a choice procedure.

However, under certain circumstances, the goal status of the food does not alter *immediately*. For example, if the food is devalued by isotonic LiCl injection following a meal, rats do not work less for the food until they have had the opportunity to *re-experience* the food by consuming it (Balleine & Dickinson, 1991). This implies that there are two representations of the food's value. After the injection, something in the rat has learned from the conditioning experience, and will react 'aversively' to the food when it is next eaten, but the *instrumental incentive value* (the value governing goal-directed instrumental action) has not yet changed. Dickinson and colleagues refer to the process by which instrumental incentive value is updated as *incentive learning* (Dickinson & Balleine, 1994).

Hedonic assessment: an implied, distinct valuation process

The system that reacts immediately (but covertly) to food devaluation procedures, is independent of the instrumental incentive value, and comes into play upon direct experience of the food has been termed an affective or *hedonic* system (Garcia, 1989). To restate this hypothesis: the devaluation procedure modifies the neural system responsible for hedonic experience, so that it will react with disgust rather than pleasure when the devalued foodstuff is next experienced. In the meantime, the more 'cognitive' incentive value remains high, so the animal still works for the devalued food. The next time the food is consumed, direct experience of the food leads to the disgust reaction being evoked, which re-writes the neural representation of incentive value and leads the animal to work less for the food in the future.

Although hedonic reactions may be conditioned and assessed in humans by direct questioning (e.g. Baeyens *et al*., 1990), it is not obvious that they can be assessed at all in other species. However, it has been suggested that taste reactivity patterns — the orofacial reactions of rodents to flavours introduced

into the mouth — are an index of hedonic experience in rats (Grill & Berridge, 1985), and indeed, they behave in a manner compatible with the role required by Dickinson and colleagues of their hedonic system, such as tracking motivational state directly (Berridge *et al.*, 1981; Berridge *et al.*, 1984; Berridge, 1991; Berridge & Robinson, 1998, p. 314).

Some other features of the incentive learning process are worth noting. Some treatments, such as hypertonic intraperitoneal LiCl injection (which causes somatic discomfort) change both 'value' representations at the same time (Balleine & Dickinson, 1992). Additionally, the incentive value of a food is initially independent of current motivational state (i.e. hunger, thirst); thus, when a hungry rat is trained to respond for food, and then sated before being tested in extinction, it will respond as much as another rat that is hungry on test, until it experiences directly the reduced value of the foodstuff when sated (Balleine, 1992). Following this re-experience, the incentive value will vary appropriately with the motivational state of the animal, so that the rat will subsequently work hard if it is hungry, but not if it is sated. This implies that both the 'immediate-assessment' (hedonic) system and the incentive value system have access to motivational state information.

Discriminative stimuli

A very brief review of the role of discriminative stimuli (S^D) is also in order. The basic procedure for establishing a stimulus as a positive S^D is to reinforce responding during the S^D, and to withhold reinforcement in its absence. Clearly, while such a stimulus does signal that the instrumental response–reinforcer contingency is in operation, and might 'set the occasion' for responding, it might also enter into direct stimulus–response associations, or act as a Pavlovian CS for the reinforcer (because the reinforcer is only delivered in the presence of the S^D). Indeed, it has been demonstrated that S^Ds do become associated with their reinforcer (reviewed in Colwill & Rescorla, 1988). Nevertheless, it is likely that a more complex explanation is also required: S^Ds have effects that cannot be explained in terms of Pavlovian associations (Holman & Mackintosh, 1981) and it has been shown there is a conditional relationship in which an S^D signals the operation of a particular response–reinforcer contingency (Colwill & Rescorla, 1990; Rescorla, 1990a; Rescorla, 1990c).

Stimulus–response habits

Although rats possess declarative knowledge of the consequences of their actions, this does not mean that they lack a procedural stimulus–response 'habit' system. There have been a number of demonstrations in which reinforcer devaluation failed to affect instrumental responding (reviewed by Adams, 1982). Investigating the reason for these findings, Adams (1982) established that overtraining is one critical determinant of whether an instrumental response becomes 'autonomous' and resistant to devaluation. Following limited experience of instrumental training (such as training in which 100 reinforcers were earned), rats' actions remained under the control of the instrumental contingency, and were responsive to reinforcer devaluation. With extended experience of instrumental responding (such as training in which 500 reinforcers were earned), their actions became habitual and resistant to devaluation (see also Dickinson *et al.*, 1995). With ratio schedules, the number of reinforcers is more relevant than the number of times the action is practised (Adams, 1982).

The schedule of reinforcement is one other factor that has an important influence on habit development. Actions trained on interval schedules are more likely to become habitual than those trained on ratio schedules (Dickinson *et al.*, 1983), presumably because of the weaker response–reinforcer contingency that such a schedule involves (Dickinson, 1994). It has been argued that a low level of experience of this contingency is the central factor governing habit development (Dickinson, 1985).

Pavlovian to instrumental transfer

Pavlovian stimuli can modulate instrumental performance by at least two mechanisms (Dickinson, 1994; Dickinson & Balleine, 1994). For example, a stimulus that predicts the arrival of sucrose solution will enhance lever-pressing for sucrose; this is the basic phenomenon of Pavlovian-to-instrumental transfer (PIT) (Estes, 1948; Lovibond, 1983). These stimuli may have a *general* motivating effect, so that a CS for a sucrose solution will enhance lever-pressing for sucrose — but also for dry food pellets — when the animal is thirsty. In addition, they may act selectively to potentiate actions with which they share an outcome (in this example, potentiating lever-pressing for sucrose but not for food); this is a *response-* or *outcome-specific* form of PIT.

Response-specific potentiation

A prototypical demonstration of this effect was provided by Colwill & Rescorla (1988, Experiment 3, abbreviated). Animals were trained to associate a stimulus with either pellets or sucrose solution. They were then trained separately to perform two instrumental actions, one for pellets and one for sucrose. The design is shown in Table 1.

Table 1. Specific Pavlovian–instrumental transfer; design of Colwill & Rescorla (1988) (S, stimulus).

Group	Training	Test
Pellet	S → pellet Lever-press (Lp) → pellet Chain-pull (Cp) → sucrose	S: Lp > Cp
Sucrose	S → sucrose Lp → pellet Cp → sucrose	S: Lp < Cp
Animals are hungry throughout.		

During an extinction test, the stimulus had a greater ability to potentiate the action with which it shared an outcome (see also Colwill & Motzkin, 1994). (Note that response specificity is implied whether the stimulus selectively elevated the action with which it had shared an outcome, or selectively depressed the action with which it had *not* shared an outcome, as Colwill & Rescorla actually observed.) It has been suggested that Pavlovian stimuli act by reinstating conditions that are more similar to those in which the instrumental action was trained (Trapold & Overmier, 1972), but this need not be the case. Colwill & Rescorla (1988) also compared the effects of a Pavlovian CS+ to those of a discriminative stimulus (S^D); though the transfer effect was qualitatively similar there were quantitative differences and they suggest differences between the roles of a CS and an S^D. The exact mechanism is therefore controversial (see Colwill & Motzkin, 1994), but the concept of response-specific PIT has been clearly demonstrated.

It should be noted that the potentiation is dependent upon the instrumental response and outcome, and necessarily dependent upon the Pavlovian US. The CS must call up a representation of the US sufficiently detailed to discriminate pellets from sucrose.

General potentiation of behaviour

Dickinson & Dawson (1987b, Experiment 2) showed that Pavlovian stimuli whose outcomes are relevant to the current motivational state may potentiate ongoing instrumental behaviour in a general fashion. Their design is illustrated in Table 2.

Table 2. General PIT; design of Dickinson & Dawson (1987b).

Group	Training	Test
Hungry	S1 → pellet	(split into two groups)
	S2 → 20% sucrose	Hungry: Lp(S1) > Lp(S2)
	Lp → pellet	Thirsty: Lp(S1) < Lp(S2)

Following training, all animals were tested in extinction. Half were tested hungry, and in this case presentation of the stimulus associated with pellets had the greatest effect to increase lever-pressing. However, half were tested thirsty; here, the liquid sucrose solution is more relevant to the current motivational state, and indeed the animals pressed more under the stimulus associated with sucrose — despite never having pressed a lever for sucrose. The potentiation is thus independent of the outcome of the instrumental action. It may be clearly differentiated from PIT based on the reinstatement of training conditions, for such an effect would predict Lp(S1) > Lp(S2) in all cases.

It may be argued that the thirsty rats in this study are showing specific suppression by S1, rather than general potentiation by S2. Indeed, there is evidence that stimuli predictive of food suppress lever-pressing in thirsty rats. A study by Balleine (1994) demonstrated such an effect, which was asymmetrical across hunger and thirst and which may reflect the fact that dry foods aggravate water deprivation and are thus aversive to thirsty rats, while fluid consumption does not aggravate food deprivation. However, this same study also provided clear evidence of a general PIT effect. Rats that were trained to lever-press for water whilst thirsty, and were then shifted to a state of hunger, pressed more when a CS for food pellets was presented (Fig. 2 of Balleine, 1994, group PEL, tested hungry). A general suppressive effect was also seen when a CS for food was presented to thirsty rats previously trained to press for water.

This general potentiation has been interpreted as *conditioned motivation* (see Dickinson, 1994): motivation is conditioned to the Pavlovian CS during training. Note that the potentiation is independent of the instrumental outcome (a stimulus paired with sucrose solution can potentiate responding for pellets), and obviously independent of the instrumental response. However, it is dependent upon the Pavlovian US because the CS only potentiates behaviour when its US is currently relevant, and therefore the US must be distinguished from others that are not.

This process is responsible for the 'irrelevant incentive' effect, in which stimuli (including contextual stimuli) associated with motivationally relevant outcomes potentiate behaviour, even in the absence of prior experience of the outcome in that motivational state (Dickinson, 1986; Dickinson & Dawson, 1987a; 1987b). That is, the level of Pavlovian–instrumental transfer depends directly on the relevance of the outcome to the current motivational state; therefore, the neural system responsible for PIT must have access to motivational state information. As the potentiation is independent of instrumental outcome, by implication it cannot affect choice behaviour; this has been demonstrated for the irrelevant incentive effect (Dickinson, 1986).

The implication is that information about the identity of the US must be available on test. In Dickinson & Dawson's (1987b) study, it would not be sufficient to have learned affective values during training (S1 → nice, S2 → nice), even if they had slight differences (S1 → superb, S2 → OK) because this could not explain the bidirectional nature of the stimulus control on test. Furthermore, there is no opportunity to learn a conditional affective value (S1 + hunger → nice, S1 + thirst → nasty, etc.) because the reinforcer has never been experienced in the motivational state present on test. Therefore, a more detailed representation of the US must be available (such as its liquidity), which can then be assessed for motivational relevance at the moment of test. The degree of specificity can be gauged by comparing similar rein-

forcers; for example, this procedure requires the discrimination of sucrose solution and pellets during the test using only conditioned cues, and rats are clearly capable of this.

As specific and general PIT have never been doubly dissociated, it remains possible that they share a common mechanism — for example, that one process provides 'vigour', which is then 'directed' to appropriate responses if they are available. However, it seems likely that PIT is separable from the 'hedonic' system discussed above: for example, PIT (in the form of the irrelevant incentive effect) can occur without updating instrumental incentive values (Dickinson, 1986).

Simple transfer tasks

Pavlovian–instrumental transfer has more often been assessed in a simpler task. The general design is illustrated in Table 3.

Table 3. Simple designs for demonstrating PIT. (ISI, interstimulus interval.)

Study	Training	Test
Lovibond (1983) Estes (1948)	S+ → pellet Lp → pellet	Lp(S+) > Lp(ISI)
Dickinson *et al.* (2000) Hall *et al.* (1999)	S+ → pellet S– → Lp → pellet	Lp(S+) > Lp(ISI) Lp(S+) > Lp(S–)
Animals are hungry throughout.		

This type of task does not differentiate between the two transfer mechanisms, as the stimulus shares a motivationally relevant outcome with the only action being tested.

Simple PIT was further investigated by Lovibond (1983), who found that Pavlovian CSs had different effects on responding under ratio and interval schedules. Under variable ratio (VR) schedules, PIT was only observed when the pre-CS response rates were very low, and the potentiation of responding was long-lasting. Essentially, the CS *restarted* subjects that had ceased responding. In contrast, when variable interval (VI) schedules were used, presentation of the CS resulted in an elevation of responding that lasted only for the duration of the CS, and the CS rate of responding was approximately proportional to the pre-CS rate of responding. This result may be interpreted in several ways. Lovibond, for example, argued that the CS was effective when there was least 'stimulus support for responding' (i.e. when temporal cues from the schedule or the animal's own responding predicted food least; Lovibond, 1983). As interval schedules engender habitual responding more readily than ratio schedules (Dickinson, 1985; 1994), PIT may reflect a process that enhances the habitual component of instrumental responding. Alternatively, PIT may boost goal-directed (non-habitual) responding by acting as a form of reminder, thus restarting responding under ratio schedules at moments when the contribution of goal-directed responding was low, and consistently enhancing responding under interval schedules, when the contribution of goal-directed responding is always low (A. Dickinson, personal communication, 1999). It is also possible that general and response-specific PIT have different effects in this regard; all these questions remain unanswered. Regardless, interval schedules appear to provide the best conditions under which to observe a stable and reliable PIT effect.

Summary

At least six processes are implicated in instrumental performance in rats. This picture, undoubtedly complex, is summarized in Figure 5. Although some interactions between Pavlovian and instrumental behaviour have already been discussed, there remain many uncertainties, such as whether central 'affect' states

(inferred from Pavlovian studies) are the same as, or interact with, one of the 'value' systems inferred to control instrumental responding.

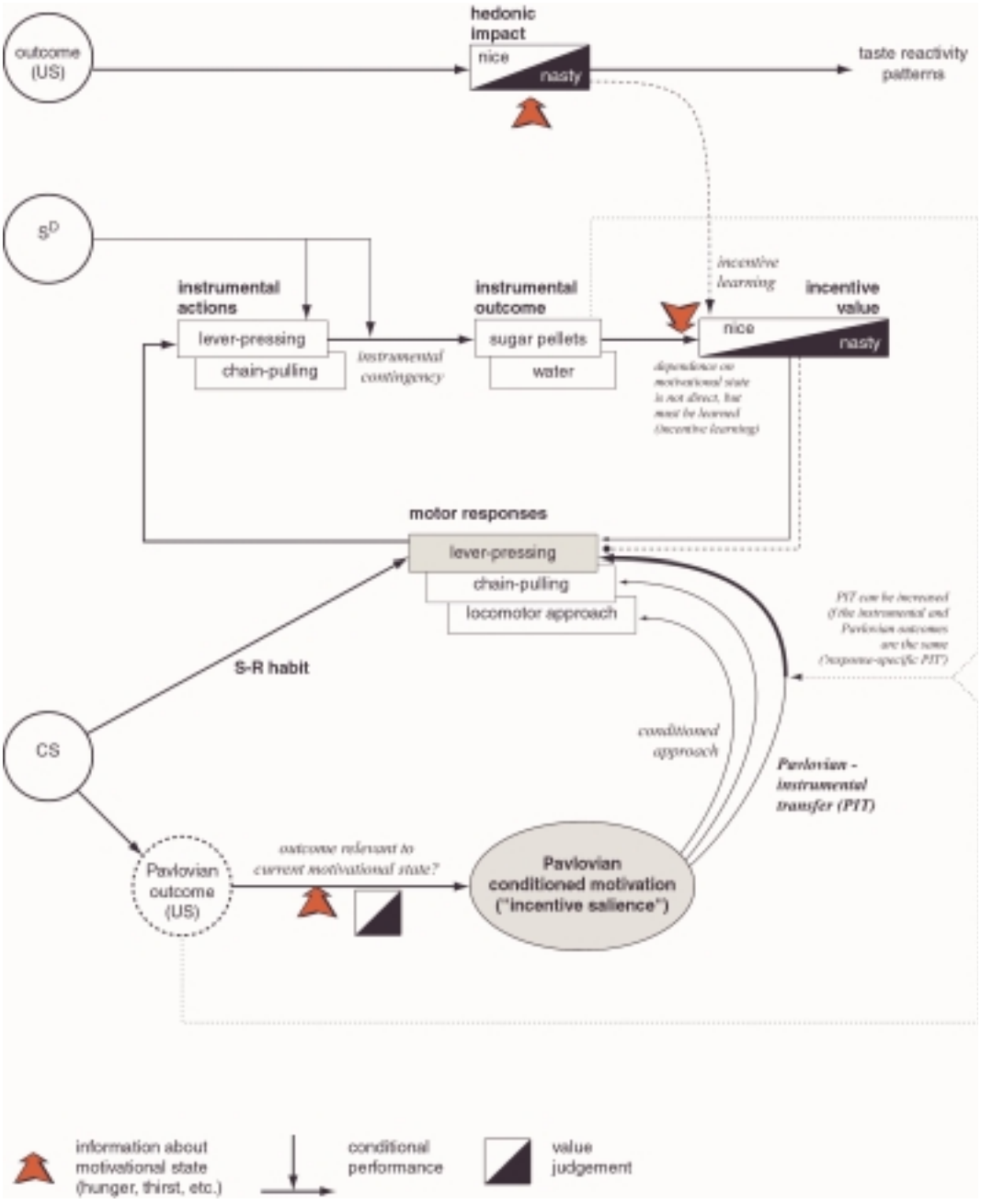

Figure 5. Some processes that contribute to instrumental behaviour in rats. An action such as lever-pressing is capable of being detected and represented in a system that can encode the contingency between this action and outcomes. When this representation is combined with a favourable representation of the instrumental incentive value of the outcome, lever-pressing is promoted. The instrumental contingencies currently in force can be signalled by instrumental discriminative stimuli (S^D). The value governing goal-directed responding is learned through direct experience of the outcome in particular motivational states; it can therefore be distinguished from a 'hedonic', or immediate-assessment value system (see text). A separate contribution to response output comes from direct stimulus–response associations (S–R habits), which can be formed through repeated training. In addition to these processes, Pavlovian conditioned stimuli (CSs) that signal a motivationally relevant outcome can enhance responding, both by providing a 'motivational boost' and by potentiating responses that share an outcome with the Pavlovian CS.

CONDITIONED REINFORCEMENT

When a CS is paired with reward, it may gain some of the properties of that reward and be capable of reinforcing behaviour itself; it is then termed a conditioned reinforcer (see Mackintosh, 1974, pp. 233–259). Stimuli are established as conditioned reinforcers by Pavlovian association, and affect instrumental performance, although it is not absolutely clear how conditioned reinforcement relates to the Pavlovian and instrumental processes described so far.

Validity of conditioned reinforcement as a concept

The concept that stimuli may themselves become reinforcing has not always been accepted (recently reviewed by Williams, 1994a). Though recent demonstrations have proven that conditioned reinforcement is a genuine phenomenon, it is sometimes difficult to interpret the role of putative conditioned reinforcers in behavioural experiments. There are two common reasons for this: one relates to the techniques used to *measure* conditioned reinforcement; and the other to the methods used to *establish* stimuli as conditioned reinforcers, as some of these methods may endow the stimuli with other functions.

It is frequently suggested that when a response-contingent stimulus predicts reward, the stimulus becomes established as a conditioned reinforcer (CRf). For example, rats can be trained under second-order schedules of reinforcement; in a typical schedule, denoted FR10(FR5:S), every fifth response produces a stimulus and every tenth stimulus is accompanied by reinforcement. In this situation, omission of the stimulus impairs performance (e.g. Arroyo *et al*., 1998). Though the stimulus is certainly paired with reinforcement, it is clear that it could play a discriminative role in this task (signalling that responding is likely to lead to reinforcement); depression of responding when the stimulus is omitted would not be direct evidence that the stimulus has conditioned reinforcing properties.

However, conditioned reinforcers have effects beyond those of an S^D. It has been shown that performance under second-order schedules depends to some extent on the association of the stimulus with primary reinforcement (reviewed by Mackintosh, 1974, p. 241). More directly, Williams & Dunn (1991) demonstrated using a choice schedule that pigeons preferred a key associated with a CRf despite this leading to more unreinforced trials. In this study, preference tracked the frequency of extinction trials in which the conditioned reinforcer was presented, unconfounded by differences in primary reinforcement or by changes in the value of the CRf itself.

Further alternative explanations have been offered for the effects of response-contingent stimuli predictive of reward; for example, it has been suggested that they exert their effects through ‘marking’ the fact that a response has been made, or ‘bridging’ delays. Although these effects undoubtedly occur, it has been shown that conditioned reinforcers have *reinforcing* properties over and above their other functions (Williams, 1991).

When conditioned reinforcement must be demonstrated directly, the cleanest technique uses a different approach, namely the ‘acquisition of a new response’ procedure (Mackintosh, 1974, pp. 235–237). In an initial phase, a CS is paired with reward; in a second phase, the CS is provided contingently upon a response that is new to the subject. If changes in general activity are controlled for by providing two such new responses (i.e. a choice procedure) and the subjects work for the CS more than they work for an unpaired stimulus, it can be concluded that the stimulus is functioning as a conditioned reinforcer.

Relationship of conditioned reinforcement to other processes controlling instrumental performance

Although it seems reasonably clear that conditioned reinforcers acquire some of the properties of the primary reinforcer, rather than simply by providing information about its availability (Mackintosh, 1974, pp. 250–259), it is not immediately clear how to relate conditioned reinforcement (CRf) to the 'value systems' governing instrumental behaviour described earlier. It is relevant to consider the relationship, though, because there is considerable evidence regarding the neural basis of responding for conditioned reinforcement, which will allow comparison with the neural basis of other factors governing instrumental performance. Several possibilities exist; it should be stated at the outset that there is relatively little experimental evidence concerning any of them.

Reinforcement of S–R habits by the CRf. One simple possibility is that the CRf reinforces behaviour in the way that Thorndike (1911) originally envisaged reinforcement in his Law of Effect — by 'stamping in' stimulus–response associations. Were this to be the case, the resulting behaviour should be resistant to devaluation of the CRf (assessable, perhaps, by Pavlovian CRs evoked by the CRf). Although there have been no direct studies of this question, extinction of a CRf rapidly affects preference even when the contingency between responding and primary reinforcement is held constant (Dunn *et al*., 1987), suggesting that responding for conditioned reinforcement is not always habitual. Whether it can *become* habitual may be a very difficult question to answer: establishment of habits with primary reinforcement requires considerable training (Adams, 1982), and if conditioned reinforcement is used with an 'acquisition of new response' procedure, the CRf–US relationship will extinguish as the response is being acquired.

General PIT. We have already seen that noncontingent CSs cannot affect *choice* behaviour via the general PIT mechanism, and a mechanism that does not affect choice could not be accepted as conditioned reinforcement. This would seem to rule out general PIT as an explanation for CRf. However, when the CS is *contingent* upon one response but not another, a general potentiation of behaviour could affect choice. Pressing a lever that produces a CRf lever effectively gains access to a stimulus that 'boosts' behaviour (albeit unselectively), while pressing a control (NCRf) lever does not. If the animal's current behaviour is boosted, the CRf would tend to boost the behaviour that caused its presentation; this mechanism could favour CRf over NCRf responding. Of course, PIT could not affect responding before the first presentation of the CRf, so seems unlikely to account for results such as those of Williams & Dunn (1991), using concurrent-chain schedules (in which responding is measured for a period before the animal earns a CRf), but it is possible that PIT contributes to responding under free-operant schedules in which CRf presentation can occur during periods of responding and the CRf is presented for a comparatively long time.

Response-specific PIT. As the response-specific PIT effect increases responding for actions that share an outcome with the Pavlovian CS, it seems unlikely to affect responding for conditioned reinforcement, because the outcome of the response (which is the CRf) is not the same as the outcome of the CRf (which is the primary reinforcer). Potentiation could only occur through this mechanism if the brain could apply a less strict definition of outcome, thus: 'lever-pressing shares an outcome with the CS because lever-pressing produces the CS and therefore (by a process of inference or association) the US'. This is akin to second-order conditioning, however, so the mechanism should not be ruled out.

Instrumental incentive value of the CRf. The final possibility is perhaps the most obvious — that the CRf becomes an instrumental goal, possessing incentive value itself, which it acquires through an affective or hedonic process. Direct demonstration of this would involve training animals to respond for a conditioned reinforcer, devaluing the CRf, and assessing responding in extinction (with appropriate controls;

cf. Adams & Dickinson, 1981). Nevertheless, even in the absence of such a demonstration, it seems very likely from the results of Williams and colleagues (Dunn *et al.*, 1987; Williams & Dunn, 1991) that conditioned reinforcers do acquire incentive value; these authors have demonstrated effects of conditioned reinforcers on choice using tasks to which PIT is unlikely to contribute and in which responding is probably not habitual.

In summary, though the evidence is far from conclusive, it seems most likely that conditioned reinforcers control behaviour by acquiring instrumental incentive value, though they might also act to potentiate responding further via PIT.

NEURAL DISSOCIATIONS WITHIN ASSOCIATIVE LEARNING

Overview of the limbic corticostriatal circuits considered in this thesis

While investigating hypothalamic function, Hetherington & Ramsay (1939) and Anand & Brobeck (1951) found that electrolytic lesions of the lateral hypothalamus (LH) appear to leave animals demotivated — with impairments in unlearned behaviour (subjects were aphagic and adipsic, with reduced sexual, exploratory, and maternal behaviour) and in learned behaviour (impaired instrumental responding). However, such lesions also disrupt the medial forebrain bundle, a fibre tract that passes through the lateral hypothalamus and contains the dopaminergic projection from midbrain dopamine (DA) neurons (the substantia nigra pars compacta, SNc, and the ventral tegmental area, VTA) to the forebrain. As lesions of this projection using the dopamine-depleting toxin 6-hydroxydopamine (6-OHDA) produced a similar pattern of behavioural impairment (Marshall & Teitelbaum, 1977), as did dopamine-depleting lesions of one its targets, the striatum (Stricker & Zigmond, 1976), attention was focused on the behavioural role of dopamine and the structures that receive dopaminergic innervation.

The basal ganglia. The basal ganglia comprise a number of subcortical nuclei, including the striatum. The striatum may be considered the 'input layer' of the basal ganglia; nearly the entire neocortex projects to it (Kemp & Powell, 1971). In turn, the striatum projects to the globus pallidus, which projects via thalamic nuclei back to the cortex; the whole makes up a 'loop'. It is a particular characteristic of basal ganglia–thalamocortical ('corticostriatal') loops that although large areas of cortex send information into the loop, only a relatively small area of cortex is targeted by the return projection. Information flow in different loops is segregated — that is, the loops operate in parallel — and the loops are named for the areas of cortex to which they project: the *motor* loop (projecting in primates to the premotor cortex, supplementary motor area, and primary motor cortex, and involved in the initiation of motor acts); the *oculomotor* loop (projecting to the frontal eye fields); the *dorsolateral prefrontal* or 'cognitive' loop; the *lateral orbitofrontal* loop, and the anterior cingulate or *limbic* loop (projecting to the anterior cingulate cortex and medial orbitofrontal cortex) (DeLong & Georgopoulos, 1981; Alexander *et al*., 1986). Indeed, functional segregation (parallel processing) is apparent even *within* each loop (see Alexander & Crutcher, 1990). The loops may also be differentiated on the basis of the parts of the basal ganglia and thalamus they pass through; thus, while inputs to the motor and 'cognitive' loops target the dorsal striatum (caudate–putamen or neostriatum), information entering the limbic loop does so through the ventral striatum. The ventral striatum consists of the nucleus accumbens (Acb), ventromedial portions of the caudate and putamen, and the olfactory tubercle; the largest component is the Acb. Within each corticostriatal loop, the basic circuitry is similar across the dorsal striatum and much of the ventral striatum (reviewed by Heimer *et al*.,

1995); it is therefore likely that the various basal ganglia loops process information in qualitatively similar ways, with the nature of the cortical target determining the apparent function of each loop.

Information processing in the basal ganglia is complex, involving not only a 'direct' pathway from striatum to globus pallidus (more specifically, to the internal segment of the globus pallidus and the substantia nigra pars reticulata) but a functionally antagonistic 'indirect' pathway from the striatum to the globus pallidus (external segment), which projects to the subthalamic nucleus, and thence to the globus pallidus (internal segment) (see Alexander & Crutcher, 1990). Cellular activity in the striatum is regulated by dopaminergic projections from the midbrain. Again, there is anatomical specificity in the dopaminergic innervation: the dorsal striatum is innervated by the SNc while the ventral striatum receives its projections from the VTA. In a further subdivision of the dorsal striatum, histochemically distinct *patches* or striosomes may be defined, which may project back to midbrain dopaminergic and cholinergic neurons, while the *matrix* circuitry is as described above (Grove *et al*., 1986; Jiménez-Castellanos & Graybiel, 1989; Gerfen, 1992b; Gerfen, 1992a; Fallon & Loughlin, 1995), though it is not clear that this distinction applies to the ventral striatum (see Heimer *et al.*, 1995). In addition, there are significant dopamine projections to cortical structures that provide information to, and receive information from, the basal ganglia, such as the prefrontal cortex and amygdala (Fallon & Loughlin, 1995).

The 'limbic loop'. This thesis will focus on the functions of the limbic loop, depicted in Figure 6. Its components include many of the structures considered part of the limbic system. The term 'limbic' was coined by Broca (1878) for the cortical structures encircling the upper brain stem (*limbus,* Latin for edge or border). The 'limbic lobe' was suggested to have a role in emotional experience and expression by Papez (1937), concepts later to be elaborated by MacLean (1949; 1952; see MacLean, 1993), who introduced the expression 'limbic system' to refer to the limbic lobe and its connections with the brainstem. The limbic system is not precisely defined: as the limbic lobe was considered the neural substrate for emotions, structures whose functions have to do with motivation and emotion have since been added to the anatomical definition. A modern definition of the limbic system in primates would certainly include cingulate and orbitofrontal cortex; the hippocampal formation, parahippocampal gyrus and mammillary bodies; anterior and medial thalamic nuclei; the nucleus accumbens and ventral pallidum; the amygdala and the hypothalamus.

The nucleus accumbens. On histochemical and anatomical grounds, the nucleus accumbens may be divided into the core (AcbC), shell (AcbSh), and rostral pole (a border zone with features of the other two compartments) (Zaborszky *et al*., 1985; Zahm & Brog, 1992). The pattern of innervation of these structures differs: in terms of connectivity, Acb may be considered as having two broad functional divisions (Brog *et al*., 1993): (1) the core, rostral pole and lateral shell; and (2) the medial shell and septal pole. Of these, the core division more closely resembles the dorsal striatum, projecting predominantly to the ventral pallidum, while the shell division also projects to subcortical structures, such as the lateral hypothalamus and periaqueductal grey, involved in the control of innate behaviours. The connections of the Acb are summarized in Table 4 and Table 5 (see Berendse *et al*., 1992; Brog *et al.*, 1993).

As a recipient of information from a considerable array of limbic structures that projects additionally to nuclei known to be involved in behavioural expression, the Acb has been suggested to represent a 'limbic–motor interface' (Mogenson *et al*., 1980). However, much of the function of the Acb is presumably related to its influence over cortical structures, the function of which are themselves somewhat mysterious. In particular, the functions of the anterior cingulate cortex (ACC) will be considered in this thesis, and the comparative anatomy of this region will be discussed in Chapter 3.

Table 4. Some inputs to the nucleus accumbens (from Brog *et al.*, 1993). Subcortical connections are nearly all reciprocal.

Region in Acb	Cortical afferents	Subcortical afferents
To all / most of the nucleus accumbens	orbital cortex posterior agranular insular cortex entorhinal cortex basal amygdala hippocampal formation (via subiculum) (Note that none of these inputs is a primary or secondary sensory area or relay.)	raphé nuclei ventral tegmental area thalamic nuclei (see Brog *et al.*, 1993 for discussion)
Shell-preferential (meaning medial shell and septal pole)	dorsal peduncular cortex infralimbic cortex pyriform cortex ventral subiculum	bed nucleus of the stria terminalis hypothalamus medial amygdala lateral habenula laterodorsal tegmental nucleus sublenticular substantia innominata lateral septal nucleus locus coeruleus
Core- or rostral pole-preferential	anterior cingulate cortex medial precentral cortex dorsal and ventral prelimbic area agranular insular cortex perirhinal cortex dorsal subiculum	dorsolateral ventral pallidum subthalamic nucleus globus pallidus substantia nigra pars compacta

Table 5. Some outputs from the nucleus accumbens (for references, see Pennartz *et al*., 1994).

Region in Acb	Efferent connections
Core	ventral pallidum subthalamic nucleus substantia nigra pars reticulata
Shell	ventral pallidum ventral tegmental area substantia nigra pars compacta hypothalamus (preoptic, medial, lateral areas) lateral septum bed nucleus of the stria terminalis lateral habenula periaqueductal grey
Indirect, via ventral pallidum	mediodorsal thalamus pedunculopontine area (part of the mesencephalic locomotor region)

Interpretation of lesion studies

Although correlative techniques such as electrophysiology and functional neuroimaging allow the functioning of the normal brain to be measured, interventional techniques (such as lesion studies or drug infusions) are required to establish a causal link between a neural structure and an aspect of behaviour. In such studies, the anatomical specificity of the method is important. The use of aspirative or radiofrequency lesions, or local anaesthetic inactivation, will destroy or inactivate neurons in the target area, but will also affect fibres (axons) passing through the target structure, potentially affecting the function of neurons whose cell bodies are elsewhere. In the present thesis, excitotoxic lesion techniques and intracerebral drug infusions are used, both of which can affect neurons in the target site selectively. Excitotoxins typically activate NMDA-type glutamate receptors on neurons, leading to abnormal Ca^{2+} influx and cell death via apoptosis or excitotoxic necrosis; reviews have been provided by Choi (1988; 1995). Table 6 shows the conclusions that may be drawn from some of these interventional techniques.

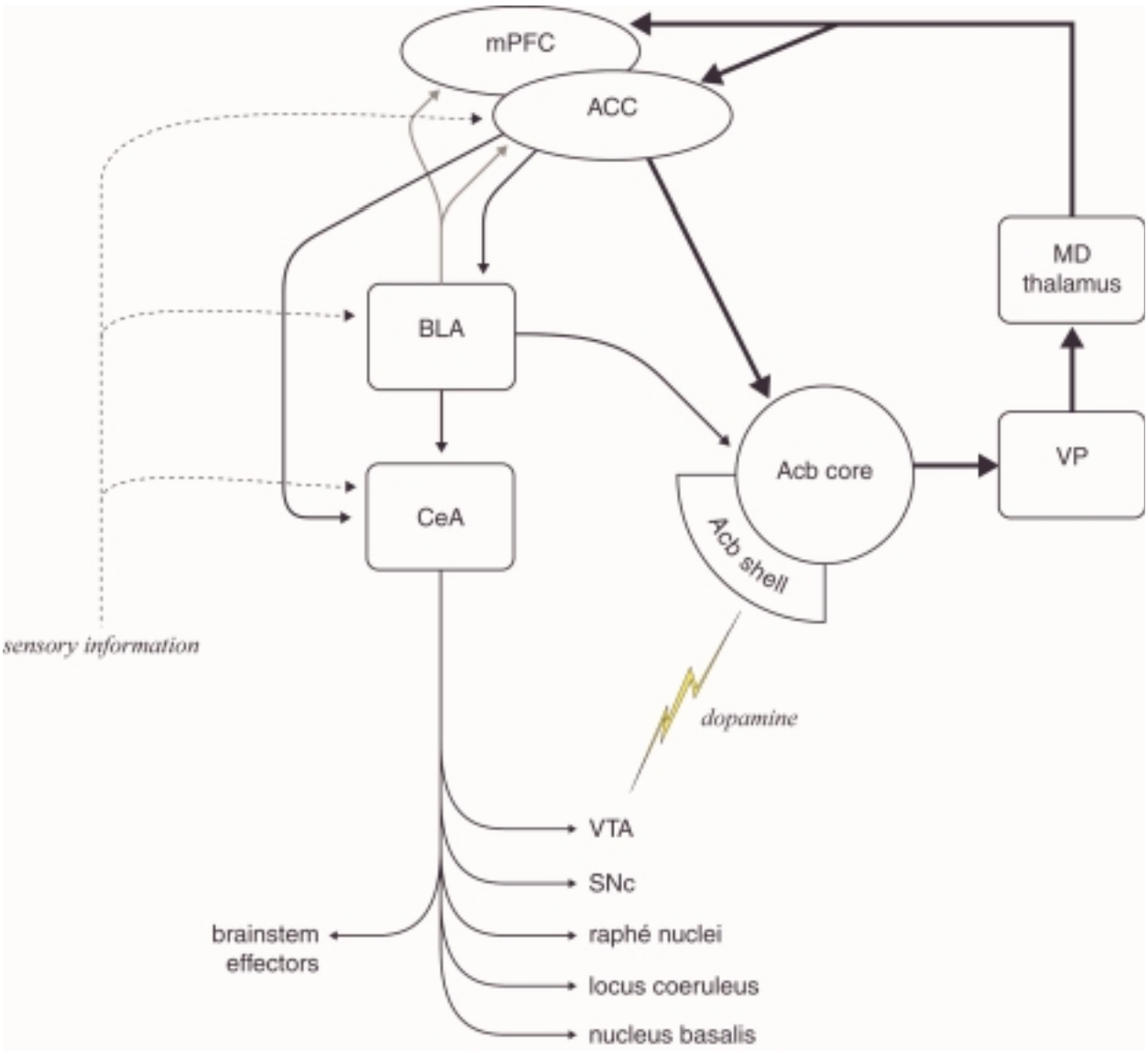

Figure 6. Part of the limbic corticostriatal loop, with associated structures. The main loop is shown in bold, together with amygdalar structures that contribute to its function in the context of appetitive approach behaviour, conditioned reinforcement and its potentiation by psychomotor stimulant drugs. For clarity, hippocampal structures are not shown. As will be discussed in the text, a functional connection between the anterior cingulate cortex (ACC) and the nucleus accumbens (Acb) core is necessary for discriminated Pavlovian approach behaviour, while the basolateral amygdala (BLA) is critical for conditioned reinforcement. The Acb is also required for the potentiation of ongoing instrumental behaviour by Pavlovian conditioned stimuli (Pavlovian–instrumental transfer; PIT). Pavlovian approach behaviour, the potentiation of conditioned reinforcement by psychostimulants, and PIT all require the central nucleus of the amygdala (CeA) and the dopaminergic innervation of the accumbens (evidence on this last point is incomplete for PIT). The integrity of the medial prefrontal or prelimbic cortex (mPFC) may be required for the perception of instrumental contingencies, and the heavy projection from the BLA to these areas of prefrontal cortex may contribute to the process by which instrumental actions are directed towards appropriate goals.

Table 6. Interpretation of lesion studies.

Manipulation	Conclusions that may be drawn from impairment	Conclusions that may be drawn from normal behaviour
Lesion, then train / test	Structure is required for learning or performance of the task	Structure is not required for learning or performance of the task, though it may still be involved
Train, lesion, test	Structure is required for performance of the task. Does not distinguish 'mnemonic' from 'motor' function.	Structure not required for performance of the task
Train in the presence of reversible inactivation; test subsequently	Either of: (a) The structure is required for task performance, and successful performance is required as part of the learning process (e.g. instrumental behaviour); (b) Learning can occur even though performance is blocked (e.g. Pavlovian conditioning), and the structure is involved in learning the task	Structure not required to learn the task
Disconnection lesion (unilateral lesion of site A and unilateral lesion of site B in the opposite hemisphere)	Site A or B must be intact bilaterally for task performance (control procedures should address this issue), or a functional connection between structures A and B is necessary for the task.	Either of: (a) A direct or indirect connection exists between the remaining A and B sites (b) Functional communication between A and B is not necessary for the task

A summary of candidate neural structures involved in Pavlovian conditioning

The focus of this thesis is firmly on the limbic structures depicted in Figure 6 (p. 35). However, to set these in a broader context, a brief review will also be provided of the role of certain other brain regions that are critical to Pavlovian conditioning.

Stimulus representation: sensory thalamus, primary and higher-order sensory cortices

The majority of sensory information concerning objects in the world reaches the brain via the sensory nuclei of the thalamus. These nuclei project directly to the amygdala, which may represent a simple, rapid route of information transfer to a structure that assesses its importance (LeDoux, 2000). But the major projections of the thalamic sensory nuclei are to primary sensory cortices, where complex attributes of stimuli are analysed. Clearly, these systems must provide CS information to Pavlovian conditioning processes; one would thus expect stimulus–stimulus associations to be organized via high-order sensory cortices. For example, cross-modal sensory preconditioning can be impaired by lesions of cross-modal sensory cortex (perirhinal cortex; Nicholson & Freeman, 2000), and following CS–food pairing, mnemonic retrieval of specific sensory aspects of the food US may depend on gustatory neocortex (see Holland, 1998). Aspects of spatial navigation, episodic memory, contextual conditioning, and other tasks requiring the integration of multiple stimuli into a 'scene' are sensitive to hippocampal lesions (e.g. Gaffan & Harrison, 1989; Selden *et al*., 1991; Gaffan, 1992; Phillips & LeDoux, 1992; Maren, 1999). Pavlovian conditioning can also lead to increases in the cortical representation of significant CSs (Weinberger, 1995; 1998a; 1998b).

The amygdala

Review

The amygdala has long been implicated in 'affective' Pavlovian conditioning, in both appetitive and aversive settings (Davis, 1992; Everitt & Robbins, 1992; LeDoux, 1992; Holland, 1997; Everitt *et al*., 2000a; LeDoux, 2000). Two of its major components are the basolateral nuclear group (BLA) (including the basal and lateral nuclei) and the central nucleus (CeA) (for anatomical reviews, see Amaral *et al*., 1992; Pitkänen, 2000).

The BLA is heavily implicated in Pavlovian fear conditioning. In a typical fear conditioning experiment, an auditory or visual CS is paired with electric shock to the feet; a CR rapidly develops in which the rat freezes for the duration of the CS (a species-specific defence reaction), while a number of autonomic changes occur. Lesions of the BLA dramatically impair the conditioned freezing response. The BLA receives information about CSs from sensory thalamic nuclei and cortices, and information about the painful US from somatosensory thalamus and cortex, while synaptic plasticity has been demonstrated in the BLA during conditioning and is necessary for conditioning to occur (for review, see LeDoux, 2000). In turn, the BLA sends a heavy projection to the CeA, which projects to a wide array of hypothalamic and brainstem structures, including the chemically-defined projection systems of the isodendritic core such as the noradrenergic locus coeruleus, the dopaminergic SNc and VTA, the serotonergic raphé nuclei and the cholinergic nucleus basalis magnocellularis (of Meynert; NBM) (Amaral *et al.*, 1992; Davis, 1992). Through these and other targets, the CeA can command autonomic responses, trigger simple skeletomotor acts such as the rat's freezing response, and regulate attentional and arousal function. It is undoubtedly the case that the BLA, which does not project directly to these structures, uses the CeA for this purpose. Thus, the dominant model of conditioning in the amygdala (Davis, 1992; LeDoux, 1992; LeDoux, 2000) suggests that CS–US association occurs in the BLA, which then expresses CRs through the CeA.

However, this picture is incomplete. Firstly, the BLA has independent projections to the ventral striatum and prefrontal (particularly orbitofrontal) cortex, giving it access to more complex response mechanisms (Everitt & Robbins, 1992; Everitt *et al.*, 1999). Secondly, the CeA also receives polymodal sensory information from association cortex (McDonald, 1998), giving it the potential to form CS–US associations, and a number of studies have demonstrated double dissociations between the effects of BLA and CeA lesions on measures of conditioning, both appetitive and aversive.

The first such demonstration concerned the modulation of instrumental behaviour. If rats are trained to respond on two concurrent schedules of food reinforcement, and one schedule additionally produces a CS terminating in mild footshock, they learn to bias their responding away from the punished lever, but also exhibit conditioned suppression when the CS occurs. Killcross *et al.* (1997b) demonstrated that BLA lesions impaired rats' ability to direct their instrumental behaviour away from the punished lever, leaving conditioned suppression intact, while CeA lesions had exactly the opposite effect (preserved instrumental avoidance with abolished conditioned suppression). A similar double dissociation using an appetitive version of the task was recently reported (Killcross *et al.*, 1998). From demonstrations such as these, a new view of amygdala function has emerged; the data are reviewed in full by Everitt *et al.* (2000a) and a summary is presented below.

The basolateral amygdala: retrieval of the current value of the US

A great deal of evidence has accumulated showing that rats with BLA lesions can acquire first-order conditioned responses, but that these responses are insensitive to reinforcer revaluation. For example, rats with BLA lesions have been shown to acquire normal conditioned responding to a CS paired with food (the CR being approach to the cup into which food was delivered). BLA-lesioned rats also showed normal acquisition of an aversion to that food when it was subsequently paired with LiCl, but failed spontaneously to adjust their responding (orienting and food cup approach) to the CS after the food was devalued (Hatfield *et al.*, 1996). Similar results have been observed in monkeys (Málková *et al.*, 1997). The most parsimonious explanation is that the conditioned responses learned by the BLA-lesioned rats were a result of direct associations between the CS and the response. They lacked the ability to use the CS to access the

value of a specific US and use that representation to alter their response. Holland (1998) defines this ability as 'mediated performance': the ability to respond based on a CS-activated representation of the US.

The idea that BLA-lesioned animals cannot use a CS to gain access to the current value of its specific US has great explanatory power. Second-order conditioning requires that the second-order stimulus becomes associated with the affective value that is called up by the first-order CS (as discussed earlier, p. 21) (see also Gewirtz & Davis, 1998): BLA-lesioned rats cannot acquire second-order conditioning (Hatfield *et al.*, 1996), cannot acquire responding under second-order instrumental schedules (Everitt *et al.*, 1989; Whitelaw *et al.*, 1996), and cannot use a first-order CS as a conditioned reinforcer (Cador *et al.*, 1989; Burns *et al.*, 1993). Clearly, the responses that still occur to the first-order CS do not support second-order conditioning, while the effects on reward devaluation (Hatfield *et al.*, 1996) demonstrate that the deficit in BLA-lesioned animals is not restricted to second-order conditioning *per se*. Specific modulation of instrumental choice behaviour by a CS also requires that the subject utilizes the motivational value of a particular US; this capability, too, depends upon the BLA (Killcross *et al.*, 1997b; Killcross *et al.*, 1998).

The formation of an association between a CS and the affective value of a US also accounts for responses such as conditioned freezing, which cannot readily be accounted for in terms of a CS–UR association. Thus, the conditioned freezing response does not resemble the UR to shock, which is characterized by agitation, jumping, vocalization and escape, but instead represents an adaptive response to danger. At the time of conditioning, therefore, there is no freezing response occurring to which a CS–UR association can be formed (see Wagner, 1970, p. 154, for discussion of this). In addition, freezing is a US-specific conditioned response: while freezing occurs to a CS for shock, it does not occur to a CS for the omission of expected food, even though both signal aversive events (as discussed above, p. 21). It seems plausible to suggest, therefore, that the BLA is critical for the acquisition of conditioned freezing because it subserves the formation of an association between the CS and a neural representation of the affective properties of the US (Bolles & Fanselow, 1980). Similarly, fear-potentiated startle may reflect the potentiation of a reflexive startle response by an affective representation retrieved by the CS, and is thereby sensitive to BLA lesions (Davis, 1997; Walker & Davis, 1997).

The central nucleus of the amygdala: stimulus–response associator and controller of the brainstem

Even though it receives neuronal afferents appropriate to support them, there is no direct evidence to suggest that the CeA is itself a site of CS–US associations; it might receive an already-associated input. However, it is clear that animals lacking a BLA can form some kinds of association, the conditioned expression of which is sensitive to CeA, but not BLA, lesions (Gallagher & Holland, 1994; Killcross *et al.*, 1997b; Hall *et al.*, 1999; Parkinson *et al.*, 2000b). The simplest analysis at present seems to be that the CeA does form simple CS–UR ('sensorimotor') associations, which do not depend upon a specific US: that is, they are independent of the identity and current motivational value of the US and are also unable to support second-order conditioning. We have suggested (Everitt *et al.*, 2000a) that the responses subserved by CeA-dependent associations especially include the modulation of reflexes organized within the brainstem, including some that might conventionally be regarded as 'affective', including conditioned suppression, conditioned orienting, and Pavlovian–instrumental transfer. These are all disrupted by CeA but not BLA lesions. Responses such as conditioned suppression may influence instrumental behaviour non-specifically, but are insufficient to modulate instrumental behaviour differentially, as assessed in choice tasks (Killcross *et al.*, 1997b).

Additionally, Gallagher, Holland and co-workers have also shown that through its projections to the reticular formation, the CeA is involved in the control of attentional aspects of stimulus processing. The

CeA plays a role in visuospatial attention during continuous-performance tasks (Holland *et al*., 2000), and also appears to regulate the *associability* of stimuli under certain circumstances (Gallagher & Holland, 1992; Gallagher & Holland, 1994; Holland & Gallagher, 1999). Associability is a learning-theory concept (e.g. Rescorla & Wagner, 1972; Pearce & Hall, 1980); it determines how much processing is devoted to a CS, and therefore indirectly determines the degree to which new things can be learned about the CS. The Pearce & Hall (1980) model of Pavlovian conditioning suggests that when a CS is reliably followed by a US, the CS may be worth responding to, but is not worth learning about: animals should confine their attention to learning about stimuli whose consequences are less well known. Associability can be increased by surprising events: for example, if a light is regularly followed by a tone, presentation of the light on its own (with the surprising absence of the tone) is predicted by the Pearce–Hall model to increase the subsequent associability of the light (e.g. Wilson *et al*., 1992; see Holland, 1997). This phenomenon — specifically, the ability to *upregulate* associability — appears to depend upon the integrity of the CeA (Holland & Gallagher, 1993b; 1993a), together with its projections to cholinergic neurons in the NBM and substantia innominata (Han *et al*., 1999), and possibly from there to the posterior parietal cortex (see Holland, 1997). Though the cellular basis of associability is unknown, it is interesting to note that Weinberger and colleagues have shown that auditory cortex receptive fields for a CS of a particular frequency expand, at the expense of other regions, when that CS is paired with an aversive US. This cortical plasticity depends upon muscarinic acetylcholine receptors and can be induced by stimulation of the NBM (see Weinberger, 1995; 1998a; 1998b). Expansion of a sensory receptive field might be one mechanism by which the associability of a stimulus could increase, as might increased attention to that stimulus directed by the attentional circuits known to exist in the posterior parietal cortex (see Posner, 1995).

Summary

This view of amygdala function is illustrated speculatively in Figure 7. When a CS predicts an appetitive US, it may form associations with sensory and motivational representations of that US (links 1 and 2 in the figure), with central affective states (3) and with unconditioned responses at some level (4). When the US is devalued, its motivational representation is in some way selectively redirected to an aversive state (not shown), so it is through link 1 or 2 that the changed response to a first-order CS occurs. It should be noted that while affective states are illustrated as ‘centres’, very little is known of the neuronal mechanism by which valence might be encoded: such information might just as easily be carried as a temporal or chemical code and be multiply represented, rather than existing in distinct spatial loci. Indeed, it has been convincingly argued that the orbitofrontal cortex (OFC), which has extensive reciprocal connections with the BLA and has also been implicated in CS retrieval of US value (Gallagher *et al*., 1999), provides an important site for the representation of affective valence (Schoenbaum *et al*., 1998; Rolls, 1999; Schoenbaum *et al*., 1999; Rolls, 2000). The exact relationship between BLA and orbitofrontal function is not clear at present; further data are reviewed when considering instrumental behaviour, below.

It is at present unclear whether the BLA is involved in representing specific sensory information about USs, required for S–S associations. Each sensory modality projects to a region of sensory cortex, a reason to question a priori whether the BLA is required, and rats can learn stimulus discrimination tasks in the absence of the BLA (Schwartzbaum, 1965; Sarter & Markowitsch, 1985; Burns *et al*., 1999). If the BLA is involved, it would therefore have to be as an ‘independent associator’ (E in Figure 2, p. 21). According to this scenario, BLA-lesioned animals make unconditioned responses and learn simple CS–UR associations, including ‘emotional’ responses, but the CS would convey no information about the identity of the US. Alternatively, the US-specific representation involving the BLA might be purely affective; in this alternative scenario, BLA-lesioned animals can learn CS–UR associations and CS–US(sensory) associa-

tions, but cannot learn CS–US(affective) associations, and the sensory representation they can activate is without affective valence (see also Holland, 1998, for a discussion of this possible dissociation). Following a recent demonstration that BLA lesions do not impair sensory preconditioning (Blundell & Killcross, 2000b), the latter interpretation seems most likely.

It is also presently unclear whether the BLA holds US-specific representations that excite general appetitive / aversive states in another structure, or itself contains this 'affective processor', or contains both. It is clearly difficult to distinguish whether BLA-lesioned animals lack affective states that may take part in associations, or merely cannot call them up via a CS; however, transreinforcer blocking and performance (but not acquisition) of second-order conditioning are two phenomena that appear to depend on simple affect, so further experiments may resolve these issues.

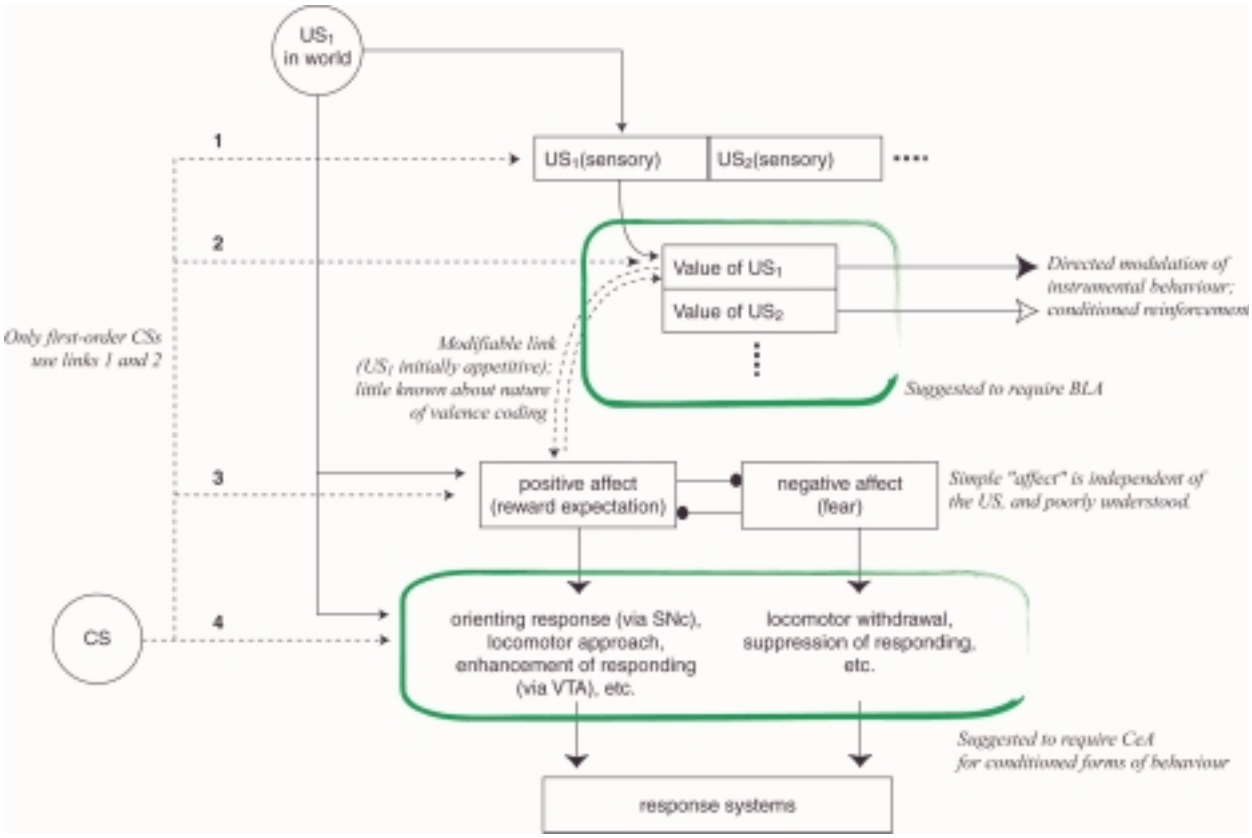

Figure 7. Schematic of representations that may be involved in Pavlovian conditioning, emphasizing the hypothesized role of amygdaloid subregions. The BLA is required for a CS to gain access to the current value of its specific US. In the figure, the CS has been associated with US_1, initially appetitive, while an unrelated US_2 maintains a separate value (connections not shown for clarity). As discussed in the text, the precise nature of the information encoded in the BLA is uncertain; here, it is illustrated as binding US-specific sensory information to an affective value. The BLA may use this information to control CeA function but also to modulate specific instrumental (choice) behaviour, as in conditioned reinforcement tasks; the Acb is a key target of this information. In contrast, the CeA is required for CS–UR learning, particularly when the response involves modulation of hypothalamic and brainstem functions. The CeA may also modulate the associability of CSs (see text), but this function is not illustrated.

Conditioning of simple CS–UR skeletal responses with high temporal precision: the cerebellum

It would be elegant if the representations encoded by amygdalar nuclei could be entirely categorized using a well-defined psychological dichotomy. It appears that we are remarkably close to this situation with the suggestion that the CeA encodes or expresses Pavlovian stimulus–response (CS–UR) associations, while the BLA encodes or retrieves the affective value of the predicted US. However, not all stimulus–response associations depend on the CeA. For example, nictitating membrane / eyeblink conditioning de-

ponse associations depend on the CeA. For example, nictitating membrane / eyeblink conditioning depends instead on the cerebellum, even though the eyeblink clearly is part of the UR to eyeshock; this circuit has been extensively mapped (see e.g. Thompson *et al*., 2000) and appears to involve CS–UR associations. Eyeblink conditioning can occur in the absence of the amygdala (even though simultaneously conditioned changes in heart rate are amygdala-dependent). In attempting to define the purview of cerebellar conditioning, Steinmetz (2000) comes to a more pragmatic, neurobiological solution: the cerebellum has been shown to be involved in associative learning when (1) a simple motor response is involved; (2) the CS–US interval is shorter than ~4 seconds; (3) the US is aversive; (4) the US not only causes a UR, but *in addition* activates the inferior olive, the 'teaching system' for such cerebellar learning. This definition fits no neat psychological category so far proposed. Applying this rationale to the amygdala, for example, would lead to the suggestion that the CeA subserves Pavlovian CS–UR associations when that response is controlled by a hypothalamic or brainstem nucleus governed by the CeA; such responses include autonomic changes, motivational arousal and attentional enhancement.

This observation has implications for general theories of learning. Belief in a general learning process has justification (Dickinson, 1980, pp. 6–9), and has led to undoubted success in describing conditioning phenomena. If associations in the cerebellum are formed according to different rules to associations in the CeA, however, there is no universal learning process. On the other hand, if such disparate systems *do* learn according to the same rules of association, why? This would imply either that highly complex associative rules are embedded on a small scale (such as at the level of the neuron) in a wide variety of neural tissue, and very consistently so, or that some as yet unknown central, cooperative learning mechanism regulates learning in widely distributed areas of the brain. There is direct psychological evidence for the latter idea (see Wagner, 1978; Dickinson, 1980, chapter 4; Baars, 1988), and the elucidation of the neural basis of this mechanism is an exciting challenge.

The anterior cingulate cortex and stimulus–reinforcer associations

One other structure within the 'limbic loop' of Figure 6 (p. 35) that has been implicated in Pavlovian conditioning is the ACC. The ACC was first implicated in aversive conditioning: it receives excitatory nociceptive information from midline and intralaminar thalamic nuclei (Hsu & Shyu, 1997), and is capable of commanding autonomic responses (Fisk & Wyss, 1997). Early studies found that aspirative lesions of the ACC, which also destroy fibres of passage, attenuated classically conditioned bradycardia in rabbits (Buchanan & Powell, 1982a). More recently, the role of the rat ACC in appetitive Pavlovian conditioning has been studied, using excitotoxic lesions (to prevent damage to fibres of passage) and the phenomenon of autoshaping.

Autoshaping. In the autoshaping task that has been studied in our laboratory, a visual stimulus (CS+) is presented on a computer screen and is followed by the delivery of food in a *different* spatial location. A second stimulus (CS–) is also presented, but never followed by food. Though the subject's behaviour has no effect on food delivery, animals develop a conditioned response in which they selectively approach the CS predictive of food before returning to the food hopper to retrieve the primary reward. Autoshaping is generally considered to be a Pavlovian conditioned response, as it can be acquired even under an omission contingency in which approach to the CS+ prevents food delivery (Williams & Williams, 1969) (see also Davey *et al*., 1981). There is an alternative possibility: that the response is instrumental and is shaped by contact with the CS, which acts as a conditioned reinforcer (see Williams, 1994a, p. 471); nevertheless, in either situation, the Pavlovian relationship between the CS and primary reward underlies autoshaping.

Using this task, Bussey *et al.* (1997a) found that ACC-lesioned rats exhibited a profound impairment in the acquisition of autoshaped responding. Lesioned rats were also impaired on an appetitive discrimination task requiring rats to learn eight stimulus–reward associations concurrently (Bussey *et al.*, 1997b). Bussey *et al.* (1996) also found that ACC lesions *facilitated* early learning of a conditional visual discrimination (CVD) task that was soluble by the formation of stimulus–response habits, but not by the formation of stimulus–reinforcer associations; Bussey *et al.* suggested that the formation of stimulus–reinforcer associations hinders performance on this task by competing with a stimulus–response system. Taken together, these studies suggest strongly that the ACC is involved in some aspect of Pavlovian stimulus–reinforcer association.

Expression of Pavlovian conditioning: amygdala–accumbens and cingulate–accumbens interactions

Though the ACC is required for the development of autoshaping, this task also requires the AcbC (though not the AcbSh) (Parkinson *et al*., 1996; Parkinson, 1998). Furthermore, AcbC lesions impair the performance of the conditioned response in rats trained before the lesion was made (Everitt *et al*., 2000b), just as they impair temporally discriminated Pavlovian approach to a single CS predictive of food (Parkinson *et al*., 1999b). Similarly, 6-OHDA-induced dopamine depletion of the whole Acb impair both the acquisition (Parkinson *et al*., submitted) and performance (Everitt *et al.*, 2000b) of autoshaping.

Indeed, the ACC is the only major limbic cortical afferent to the Acb that is required for autoshaping, as lesions of BLA, dorsal or ventral subiculum, medial prefrontal cortex (mPFC), or posterior cingulate cortex (PCC) have no effect on its acquisition (Parkinson *et al.*, 1996; Bussey *et al.*, 1997a; Parkinson *et al.*, 2000b).

It seems likely, then, that stimulus–reward associations stored or retrieved by the ACC gain behavioural expression through the Acb. This hypothesis was tested directly by Parkinson *et al*. (2000c), who used a ‘disconnection’ procedure, in which asymmetric unilateral lesions of both the ACC and Acb were made in order to prevent communication between the two structures. The disconnection impaired autoshaping, though single unilateral lesions of either structure did not.

Autoshaping is not the only form of Pavlovian conditioning in which the Acb appears to give behavioural expression to associative information arising from limbic cortical or quasi-cortical afferents. At least three other tasks have been shown to operate similarly. The first is the expression of a conditioned place preference; this depends on the BLA, but also on the Acb, and a disconnection lesion of the two structures abolishes behavioural expression (Everitt *et al*., 1991). The second is second-order conditioned approach: Setlow *et al.* (2000) recently demonstrated that BLA–Acb disconnection impairs the acquisition of second-order conditioned approach behaviour, but not second-order conditioned orienting, or first-order conditioned approach — consistent with the known involvement of the BLA in second-order conditioning (Hatfield *et al.*, 1996), and the Acb in conditioned approach (Parkinson *et al.*, 1999b; Everitt *et al.*, 2000b; Parkinson *et al.*, 2000c). The third is responding for conditioned reinforcement, discussed further below. Briefly, lesions of the BLA impair responding for conditioned reinforcement (Burns *et al.*, 1993); injection of amphetamine into the Acb dramatically enhances responding for conditioned reinforcement (Taylor & Robbins, 1984; Burns *et al.*, 1993), and this enhancement depends on the integrity of the BLA — again suggesting expression of amygdala-dependent information via the Acb.

Projections from the ACC and BLA to the Acb are direct and glutamatergic. In addition to these highly specific, information-rich projections, recent evidence suggests that the amygdala — specifically, the CeA — may also modulate Acb function via a different route. The CeA does not project to the AcbC (Zahm & Brog, 1992; Brog *et al.*, 1993; Parkinson, 1998) or the AcbSh (Zahm *et al*., 1999, pp. 1119/1124), but does project to the VTA, the source of the dopaminergic innervation of the Acb (Amaral

et al., 1992, p. 35; Fudge & Haber, 2000). It may be that the CeA is capable of regulating Acb DA; in accordance with this hypothesis, lesions of the CeA impair the acquisition of autoshaping (Parkinson *et al.*, 2000b) and the potentiation of responding for conditioned reinforcement by intra-accumbens amphetamine (discussed further below; Robledo *et al.*, 1996), tasks that depend on Acb DA (Taylor & Robbins, 1986; Parkinson *et al.*, submitted).

The anterior cingulate cortex: unanswered questions

Evidence has been summarized that the ACC is involved in Pavlovian conditioning. However, this evidence stems from a limited range of tasks, and it is unclear exactly what role the ACC plays. The pattern of impairment that Bussey *et al.* (1997a) found in autoshaping was not complete absence of responding to the CS+ or the CS–, but rather a loss of discrimination through increased CS– responding. Given the known role of the prefrontal cortex in response inhibition (summarized by Roberts *et al.*, 1998), it is possible that the ACC plays a unique inhibitory role within the 'limbic loop' in the expression of Pavlovian conditioning. Similarly, it is not known whether the ACC is critical for the expression of autoshaping, as well as its acquisition. Moreover, there are a number of other gaps in the story. It is not known whether the ACC is involved in all forms of conditioned approach, whether (given its relationship to the BLA and the Acb) it is another important contributor to the effects of conditioned reinforcers or other Pavlovian influences on instrumental performance, or whether excitotoxic lesions of the ACC impair aversive as well as appetitive conditioning. These questions will be addressed in Chapter 3, where a full review will also be made of the comparative anatomy and functions of the ACC across different species.

A summary of candidate neural structures that influence instrumental performance

Instrumental performance: outside the ventral striatum

The multifactorial view of instrumental performance outlined earlier is relatively young (see Dickinson, 1994; Dickinson & Balleine, 1994); as a result, few studies have as yet investigated the neural basis of theoretically well-defined processes contributing to instrumental performance. Of necessity, then, this summary will be brief, and will focus on the contribution of the Acb.

Contingency detection: medial prefrontal cortex and dorsal hippocampus

Demonstration that a structure is necessary for detection of action–outcome contingencies requires more than showing that an animal cannot acquire instrumental responding in its absence. Indeed, were one to prevent an animal from perceiving contingencies, there is every reason to think that instrumental performance *would* be acquired, via a habit system. Explicit tests of contingency perception are thus required. For example, rats may be trained to perform two actions concurrently for two different food rewards; in addition, one of those reinforcers may be delivered noncontingently with respect to the subjects' behaviour. The degree of action–outcome contingency for this reinforcer, P(outcome | action) – P(outcome | no action), is thus selectively degraded. In one of the few lesion studies to use this technique, Balleine & Dickinson (1998a) found that although lesions of prelimbic cortex did not prevent rats acquiring instrumental performance, or, in separate tests, from discriminating between the two actions and the two reinforcers, they rendered the rats insensitive to this contingency manipulation, suggesting that such rats might truly be 'creatures of habit'. Similar results have been obtained with lesions (albeit electrolytic) of the dorsal hippocampus (Corbit & Balleine, 2000b). However, hippocampal lesions appeared not to impair contingency learning *per se* but instead impaired sensitivity to the background, noncontin-

gent reinforcement; these effects may have resulted from a failure of contextual conditioning (Corbit & Balleine, 2000b) (see p. 36).

Indirectly, these experiments also provide an insight into habit formation, for they imply that a habit can develop without the subject ever possessing knowledge about the instrumental contingency. These studies suggest that habits 'build up' independently of goal-directed action, likely as a function of the number of reinforcers received (Adams, 1982), and are *exposed* in normal animals at the point that the action–outcome contingency no longer controls performance (Dickinson, 1985), though studies of the dysfunctioning brain cannot rule out the possibility that interaction between the two systems occurs when the intact animal learns.

Incentive value: insular cortex and other candidates — hypothalamus, ventral pallidum, amygdala, orbitofrontal cortex

Balleine & Dickinson (1998a; 2000) also investigated the role of the insular cortex, the primary gustatory cortex in the rat (Norgren, 1995), in incentive learning for food rewards. Lesioned rats performed normally on the contingency test described above. In addition, a specific satiety test was conducted, in which the rats were fed one of the foods to satiety, thus giving them the opportunity to learn that this food had reduced value in the sated state. The rats only ever experienced the other food whilst hungry. Finally, the rats' instrumental performance was tested in extinction while sated. While sham-operated control rats responded less for the reward that had been devalued, insula-lesioned rats failed to make this discrimination. However, in a further test in which the reinforcers were actually delivered, they discriminated immediately. This suggests that the insula cortex is not a critical structure for determining incentive value *per se*, but is critical for storing or retrieving the memory of the incentive value in the absence of the reward. As incentive value can be retrieved via tastes (Rescorla, 1990b; Balleine & Dickinson, 1998b), this accords with the known functions of insular cortex (see p. 36 for a similar view derived from studies of Pavlovian conditioning), although it implies some degree of dissociation between primary perception of taste and taste memory.

Incentive learning depends upon the availability of information regarding the motivational state of the animal. Are there any obvious neural candidate providers of such information? The hypothalamus, in the ventral forebrain, is such a candidate. The hypothalamus serves as the final controller of diverse bodily homeostatic systems, including endocrine function (via the pituitary gland), thermoregulation, autonomic control, and circadian rhythmicity. It plays a key role in initiating 'consummatory' behaviours, such as eating, copulation, and acts of aggression (for reviews see Swanson, 1987; Simerly, 1995). It is also the brain region responsible for detecting many of the physiological variables relevant to motivational states such as hunger and thirst; for example, it responds to blood glucose levels, gut hormones released in response to feeding, tissue osmolality, and systemic hormones released in response to fluid depletion (reviewed briefly by Kupferman, 1991). Indeed, the gut 'satiety hormone' cholecystokinin (CCK) has been shown to affect incentive learning directly (Balleine & Dickinson, 1994; Balleine *et al*., 1995a), as benzodiazepines do (Balleine *et al*., 1994).

If the incentive learning hypothesis presented above is correct, the next stage in the assignment of incentive value is hedonic experience. As discussed earlier, there are only limited techniques available for assessing the hedonic impact of foods. If the taste reactivity test (Grill & Berridge, 1985) is accepted as a measure of hedonic impact, a variety of anatomical and neurochemical systems contribute to hedonic experience, including opioid systems in the AcbSh, benzodiazepine-sensitive systems in the brainstem, and a possible common pathway in the ventral pallidum (reviewed by Berridge, 1996; Berridge & Robinson,

1998, pp. 316–317), though dopaminergic systems appear not to play a role in hedonic experience (Berridge & Robinson, 1998).

Two other major structures have been implicated in the representation of value and the control of behaviour. These are the amygdala and orbitofrontal cortex.

As discussed above, both the BLA and CeA contribute to 'affective' Pavlovian responses; the BLA is suggested to be critical for a process by which a CS retrieves the affective value of a US, and for directing instrumental behaviour accordingly (Everitt *et al.*, 2000a). A prime example of such direction is responding for conditioned reinforcement, in which the BLA directs the selection of actions according to the acquired value of the conditioned reinforcer (Cador *et al.*, 1989; Burns *et al.*, 1993). Although conditioned reinforcers may have multiple attributes, it is at least reasonable to suggest that they possess incentive value (see p. 31).

Tests of contingency perception and incentive learning in BLA-lesioned rats (A. Dickinson and B.W. Balleine, unpublished observations) are consistent with the hypothesis that the BLA is involved in determining incentive value, but are not conclusive. In the specific satiety test described above, BLA-lesioned rats were impaired at discriminating between the devalued and valued reinforcers, both in an extinction test and a reinforced test (see also Málková *et al.*, 1997). However, they were also insensitive to the contingency test described above, suggesting either (1) that they could not perceive the contingency, (2) that they could not discriminate the actions, or (3) that they could not discriminate the reinforcers. BLA-lesioned rats performed normally in a different form of contingency test, using a single reinforcer, in which subjects had to perform two actions (A1 and A2) in the specific order A1→A2 in order to obtain food (a task described in Balleine *et al.*, 1995b). Just like sham-operated rats, BLA-lesioned rats selectively increased the probability of making the chained response A1→A2 compared to the three other possible response patterns (A1→A1, A2→A2, A2→A1). This suggests that they could at least discriminate the two actions, but it does not rule out deficits in reinforcer discrimination, making the BLA's specific contribution to incentive value unclear.

In this respect, the other connections of the BLA should not be neglected. For example, the connection between the BLA and the mPFC has recently been shown to be involved in the ability of rats to modulate instrumental choice behaviour in response to conditioned punishment (Coutureau *et al.*, 2000); the anatomical connection between the BLA and the mPFC (Pitkänen, 2000) might conceivably represent a functional link between incentive value and instrumental contingencies. Additionally, the BLA is extensively and reciprocally connected to the orbitofrontal cortex (reviewed by Öngür & Price, 2000), which has been widely suggested to guide behaviour based on the anticipated value of different actions (Nauta, 1971; Damasio, 1994). In primate orbitofrontal cortex, cells may be found that respond to reward but discriminate between different rewards in doing so (Schultz *et al.*, 1998; 2000). The orbitofrontal cortex is a particularly strong candidate for a representation of incentive value, as its neurons respond rapidly to changes in the reward value of specific foods. For example, when a monkey is fed to satiety with a particular food, the orbitofrontal cortex responses to its flavour or odour decline, while the responses to other foods are unaffected (see Rolls, 2000), paralleling the behavioural change induced by sensory-specific satiety. Like the amygdala, the orbitofrontal cortex is well placed to process specific value information, as it receives projections from polymodal sensory cortex in addition to motivational state information from the hypothalamus. The relationship between the orbitofrontal cortex and the amygdala is at present unclear. Although Rolls has suggested that primate orbitofrontal cortex acts as a highly flexible system that takes over functions of the more primitive amygdala (Rolls, 2000), Schoenbaum *et al.* (1999) found evidence that, in the rat, the BLA rapidly learns to respond to CSs according to the motivational value of the

US, while changes in the electrophysiological response of orbitofrontal cortex cells follow later and are more clearly related to choice behaviour. Recently, direct evidence for a functional connection between the BLA and orbitofrontal cortex has been provided by Baxter *et al.* (2000), who showed that disconnecting these two structures impaired the ability of rhesus monkeys to adjust their choice behaviour in response to reinforcer devaluation.

Stimulus–response habits: the dorsal striatum?

The dorsal striatum (a component of the 'motor loop' of the basal ganglia; Alexander *et al.*, 1986; Alexander & Crutcher, 1990; Alexander *et al.*, 1990), together with its dopaminergic innervation, facilitates stimulus–response coupling — that is, the initiation of motor actions in response to environmental stimuli (Robbins & Everitt, 1992). It is natural to suggest that such stimulus–response coupling may underlie S–R habits in instrumental behaviour. A role for the basal ganglia in habit formation was originally suggested by Mishkin *et al.* (1984), who saw a habit as a direct stimulus–response association that was learned slowly but was stable. A recent review of this concept is provided by White (1997). Much of the subsequent work on this issue has proved controversial (see Wise, 1996; Wise *et al.*, 1996), and some of the best evidence for a long term change in behaviour that is dependent on the striatum is from an experiment by Packard & McGaugh (1996). They trained rats in a T-maze with one arm consistently baited. This task is soluble by two methods: repeating the reinforced response (the physical response of turning left or turning right), or approaching the place where food was found (a 'place response'). These may be distinguished by letting the rat approach the choice point from the opposite direction. After 8 days of training, most rats made place responses. Inactivation of the dorsal hippocampus with lidocaine (lignocaine) on the test session eliminated this tendency, such that the rats showed neither place nor motor response learning, but inactivation of the dorsolateral caudate nucleus had no effect. After 16 days of training, however, most rats made the motor response that had been reinforced. Inactivation of the hippocampus had no effect, whilst inactivating the caudate eliminated 'motor' responding and reinstated place responding. Therefore, development of a stimulus to motor response mapping takes place slowly during reinforced training and comes to dominate behaviour, and its performance depends on the caudate nucleus. However, it should be noted that studies like this one do not always satisfy the definition of 'habit' used in the discussion of instrumental behaviour above. For example, few such studies have tested the effect of reinforcer devaluation on performance of the presumptive habit.

The nucleus accumbens, dopamine, and the impact of Pavlovian conditioned stimuli

Goal-directed action does not require the nucleus accumbens

The available evidence suggests that the Acb is not required for goal-directed action. Balleine & Killcross (1994) studied rats with excitotoxic lesions of the Acb performing a lever-pressing task. They established that these rats remained sensitive to a change in the instrumental contingency (from response-contingent to non-contingent reinforcer delivery); in addition, they were sensitive to a change in the value of the instrumental outcome. By the criteria of Dickinson & Balleine (1994), these rats remained capable of goal-directed action. Similarly, dopamine receptor antagonists do not affect the representation of value that governs goal-directed action (the instrumental incentive value; Dickinson *et al.*, 2000). Insofar as the issue has been addressed experimentally, stimulus–response habits also persist following Acb lesions or dopamine depletion (Robbins *et al.*, 1990a; Reading *et al.*, 1991), although these studies did not use outcome devaluation tests to demonstrate that behaviour was habitual.

At first sight, these results are inconsistent with studies showing that manipulations of Acb affect responding for food. For example, Kelley *et al*. (1997) demonstrated that NMDA receptor blockade of the AcbC impaired the acquisition of a lever-press response for food, though not its subsequent performance on an VR2 schedule. Similarly, Salamone and colleagues have shown that dopamine depletion of Acb reduces the ability of to perform instrumental responses when the work requirement is high (e.g. Aberman & Salamone, 1999). Indeed, Balleine & Killcross (1994) found that Acb-lesioned rats responded at a lower level than controls.

However, when simple reinforcement schedules are used, there are many potential influences on performance. One such influence is the impact of Pavlovian conditioned stimuli (CSs) in the environment, and, as suggested by Balleine & Killcross (1994), it is for the impact of these stimuli that the Acb appears critical. As discussed above (p. 42), the Acb is required for autoshaping, in which locomotor behaviour is controlled by appetitive Pavlovian conditioned stimuli. In addition, it has been shown to be involved in at least two situations in which Pavlovian stimuli affect instrumental behaviour.

Responding for conditioned reinforcement is affected by accumbens manipulations

Following the suggestion by Hill (1970) that an important mechanism of action of psychostimulant drugs was to enhance the effects of conditioned or secondary reinforcers, amphetamine was shown to potentiate responding for conditioned reinforcement when injected directly into the Acb (Taylor & Robbins, 1984). In the prototypical task, rats are first trained to associate a CS with the delivery of primary reinforcement. In a subsequent extinction test, they are presented with two levers; responding on the CRf lever results in delivery of the CS, while responding on the NCRf lever has no consequence. Intra-accumbens dopamine agonists greatly enhance responding for the conditioned reinforcer, an effect that is anatomically, behaviourally and pharmacologically specific (Taylor & Robbins, 1984; Taylor & Robbins, 1986; Cador *et al*., 1991).

Subsequent studies have demonstrated that the ability of amphetamine to potentiate responding for conditioned reinforcement depends on the integrity of the AcbSh (Parkinson *et al.*, 1999b), the dopamine innervation of the accumbens (Taylor & Robbins, 1986; Cador *et al.*, 1991; Wolterink *et al*., 1993), and the CeA (Robledo *et al.*, 1996), once again raising the possibility that the CeA normally plays a part in controlling Acb DA during appetitive Pavlovian tasks (see p. 43).

However, the glutamatergic inputs to the Acb involved in conditioned reinforcement appear to differ from those involved in autoshaping (for which projections from the ACC to the AcbC are critical). The efficacy of conditioned reinforcers is impaired by lesions of the BLA (Cador *et al.*, 1989; Everitt *et al.*, 1989; Burns *et al.*, 1993), but not the mPFC or the ventral subiculum (Burns *et al.*, 1993). Information of some sort about conditioned reinforcement must arrive at the Acb for its effects to be potentiated selectively by intra-Acb amphetamine — either the Acb must have *direct* information regarding the motivational significance of the CRf, or other structures that cause the animal to respond for CRf must provide the Acb with information about the identity of the current prepotent response, in order for intra-Acb amphetamine to potentiate this response selectively. Thus it appears that information regarding the conditioned value of the CS depends upon the BLA and is conveyed to the Acb (though not necessarily directly or exclusively), where its effects can be potentiated or 'gain-amplified' by dopamine (Robbins & Everitt, 1992). The BLA projects strongly to the Acb (both core and shell; Brog *et al.*, 1993), and while shell lesions abolish the effects of intra-Acb amphetamine, lesions of the core alter the normal response to intra-Acb amphetamine, such that amphetamine increases responding on both levers — a loss of response selectivity (Parkinson *et al*., 1999c). Whether the orbitofrontal cortex or ACC also contributes to responding

for conditioned reinforcement is unknown at present; the role of the ACC is investigated directly in Chapter 3.

It remains a mystery as to precisely how the core and shell subdivisions of the Acb interact in this task. Apparently, information regarding the conditioned reinforcer arrives at the core and the shell (directly or indirectly from the BLA), but the ability of amphetamine to amplify the effects of this information depends upon the dopaminergic innervation of the Acb and the integrity of the shell, while the response selectivity of this amplification depends upon the core. Perhaps the enhancement of responding induced by intra-shell amphetamine is directed by the core towards the correct response. Though the core and shell do not project to each other directly (Brog *et al.*, 1993), the shell may modify the information passing through the core via indirect routes: Haber *et al*. (2000) have shown that the shell projects to regions of the VTA that innervate the shell itself, but also to VTA regions that project to the core; thus, the shell may exert control over dopamine function in the core. (Similarly, the core may be able to exert control over the dopamine projection to itself and to the central striatum, which may control the dorsolateral striatum in an 'ascending spiral' — a progression from limbic, through cognitive, to motor corticostriatal loops; see Figure 8.) Nevertheless, if this scheme is applicable, it is unclear why shell lesions block the effect of amphetamine injections into the core (Parkinson *et al.*, 1999b). Alternatively, it may be that intra-Acb amphetamine's effects on the vigour and direction of behaviour (dependent upon the AcbSh and AcbC, respectively) are not integrated within the Acb, but are integrated at downstream sites (a possible candidate being the ventral pallidum; Fletcher *et al*., 1998).

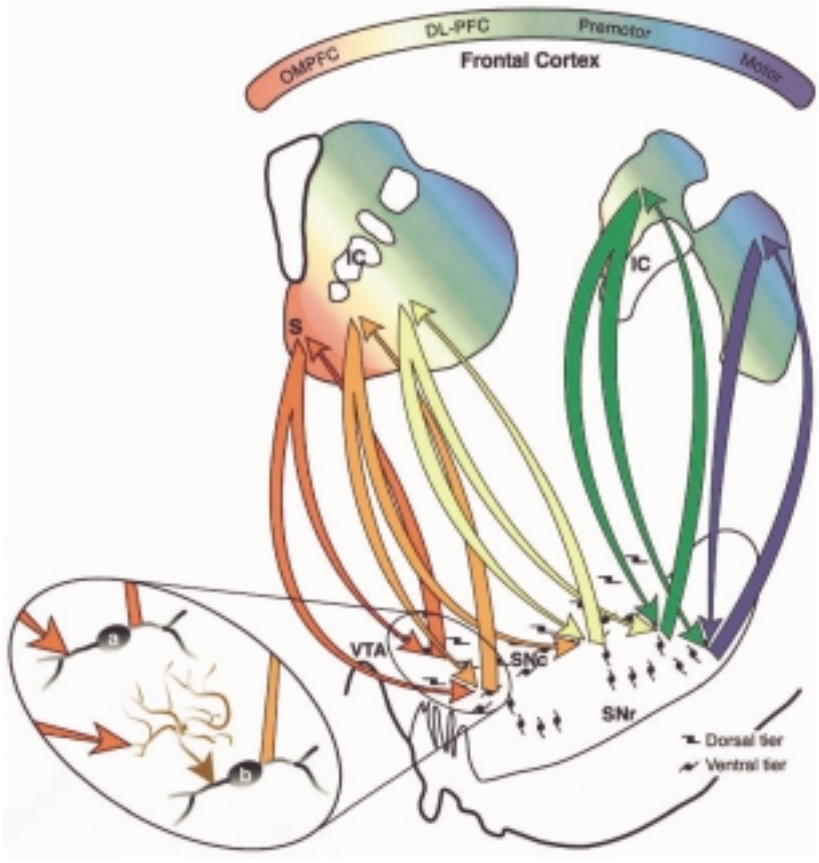

Figure 8. Organization of striatonigrostriatal projections in the primate, illustrating one putative, dopaminergic mechanism by which corticostriatal loops influence each other in a hierarchy (Haber *et al.*, 2000). The colours illustrate the corticostriatal loops (*red*, limbic; *green*, associative; *blue*, motor). The AcbSh projects to regions of the VTA that innervate the AcbSh (*red*), but also the AcbC (*orange*). Similarly, the AcbC projects to areas of the VTA that innervate itself, but also to regions that project to the dorsomedial striatum. This spiral continues through more dorsal striatal regions (*yellow* → *green* → *blue*). The magnified oval region illustrates a hypothetical regulatory mechanism: striatal projections to those VTA neurons providing a closed-loop feedback projection terminate directly on the dopaminergic cell, inhibiting VTA neuron firing; however, striatal projections to those VTA neurons providing a feedforward projection to a different striatal region terminate on inhibitory interneurons, disinhibiting the dopaminergic innervation of the adjacent region (in this case, the AcbC). *DL-PFC*, dorsolateral prefrontal cortex; *IC*, internal capsule; *OMPFC*, orbital and medial prefrontal cortex; *S*, nucleus accumbens shell; *SNc*, substantia nigra pars compacta; *SNr*, substantia nigra pars reticulata; *VTA*, ventral tegmental area. Reproduced from Haber *et al*. (2000).

Surprisingly, animals with lesions of the AcbC or AcbSh retain the basic conditioned reinforcement effect — the ability to respond preferentially on a lever delivering an appetitive CS — even though their response to psychostimulants is altered. One possibility is that the expression of conditioned reinforcement itself does not depend entirely on Pavlovian processes. Clearly, Pavlovian conditioning is the mechanism by which a stimulus is established as a conditioned reinforcer. However, the *expression* of this learning might be through several mechanisms; in particular, it must be considered that the conditioned reinforcer becomes a true declarative instrumental 'goal' (discussed above, p. 31). Given that the accumbens is not necessary for animals to represent the value of an instrumental goal (Balleine & Killcross, 1994), it might not then be expected that Acb lesions would remove all effects of a conditioned reinforcer. The BLA is known to be important for the representation of the value of conditioned reinforcers (Cador *et al.*, 1989; Burns *et al.*, 1993; Killcross *et al.*, 1998); other candidate structures for the representation of the value governing instrumental responding for CRf are the insular cortex and orbitofrontal cortex (see pp. 44–46). In contrast to the results discussed above, however, Dix *et al.* (2000) recently reported that excitotoxic lesions of the whole Acb impaired the ability of rats to direct instrumental behaviour on the basis of conditioned punishment; it remains to be seen what the effects of selective AcbC and AcbSh lesions would be on this task. While this result might be interpreted as a difference between the circuits mediating the effects of appetitive and aversive CSs, it remains possible that both the AcbC and AcbSh contribute to the basic effect of conditioned reinforcement and that excitotoxic lesions of the whole Acb would impair responding for conditioned reinforcement, though dopamine depletion does not (Taylor & Robbins, 1986).

Pavlovian–instrumental transfer is impaired by lesions involving the nucleus accumbens

Conditioned reinforcement is a phenomenon by which a Pavlovian CS is delivered contingent upon responding. However, the accumbens is also critical for the impact of *noncontingent* Pavlovian conditioned stimuli. This has been demonstrated clearly by Pavlovian-to-instrumental transfer (PIT) experiments. If an animal is trained to press a lever for food and subsequently tested in extinction, presentation of a Pavlovian CS that predicts the same food increases the rate of lever-pressing (Estes, 1948; Lovibond, 1983). Lesions of the AcbC (Hall *et al.*, 1999) abolish PIT (see also de Borchgrave, 1995), as does systemic treatment with dopamine receptor antagonists (Smith & Dickinson, 1998; Dickinson *et al.*, 2000). A recent study also demonstrated that PIT can be enhanced by intra-accumbens amphetamine in the same way that conditioned reinforcement is. Wyvell & Berridge (2000) trained rats to respond on a lever for food, and also gave them associative pairings of a lever / light CS with that food. In a subsequent extinction test, they found that intra-Acb amphetamine (targeted at the AcbSh) increased the ability of the CS to potentiate responding, whether the CS was located within the lever (in which case the results might reflect a potentiation of autoshaping rather than PIT) or was a diffuse auditory stimulus. Finally, PIT is also impaired by CeA lesions (Hall *et al.*, 1999), leading Hall *et al.* (1999) to speculate that the ability of an appetitive Pavlovian CS to potentiate instrumental behaviour depends on the mesolimbic dopamine system, presumably under the control of the CeA (see pp. 43 & 47).

As described earlier, PIT can be subdivided into a general arousing effect of appetitive Pavlovian stimuli and a more informational component by which Pavlovian CSs selectively potentiate instrumental behaviour with which they share an outcome. It remains to be seen whether the arousing (general) and informational (specific) mechanisms by which noncontingent stimuli potentiate behaviour are the same as those involved for contingent stimuli (conditioned reinforcers). In both cases, such evidence as is available suggests that the informational component is subserved by glutamatergic projections from limbic structures such as the amygdala and ACC, with that information arriving directly or indirectly at the Acb, whilst the arousal component depends upon ascending projections from the isodendritic core to the Acb

(Taylor & Robbins, 1984; Cador *et al.*, 1989; Bussey *et al.*, 1997a; Han *et al.*, 1997; Hall *et al.*, 1999). Consider a response-specific PIT task. If PIT is truly comparable to the potentiation of responding for CRf by intra-Acb amphetamine, then one would expect the AcbC to be responsible for the response selectivity of PIT, and the AcbSh to be critical for the potentiation itself. If, on the other hand, response-specific PIT has a great deal in common with 'simple' PIT, one might expect AcbC lesions to abolish PIT entirely, as in the study of Hall *et al.* (1999). These predictions will be compared in Chapter 4 by testing the effects of lesions to the nucleus accumbens core and shell on response-specific PIT.

Contribution of the nucleus accumbens to complex behaviour

The role of these motivational processes in performance under different schedules of reinforcement is imprecisely understood. From an economic point of view, there is a high probability of executing an action when the motivation to perform that action exceeds the response costs, which include the work-related costs (effort). Schedule performance depends on these two variables; indeed, the progressive-ratio (PR) schedule is based on these principles. Salamone and colleagues have demonstrated that 6-OHDA-induced dopamine depletion of Acb impairs the ability of animals to overcome response costs (Salamone, 1994). Thus, DA-depleted rats will forgo the opportunity to press a lever for a preferred food, instead consuming more of a less-preferred but freely available food (Salamone *et al.*, 1991; Cousins *et al.*, 1993). Similarly, dopamine depletion impairs responding on high-rate but not on low-rate schedules (McCullough *et al.*, 1993; Salamone *et al.*, 1993; Sokolowski & Salamone, 1998; Aberman & Salamone, 1999). Of course, some of these results may be explained in terms of motoric impairments (such as a reduction in the maximum possible rate of responding). Cousins *et al.* (1996), however, tested rats in a T-maze in which one arm led to a large reward, but was obstructed by a barrier over which the rats had to climb, while the other arm, though it led to a small reward, was unobstructed. Cousins *et al.* found that while Acb DA depletion significantly reduced rats' preference for the arm that contained the barrier in this situation, DA depletion had minimal effects on rats' ability to climb the barrier when no alternative reward was available, suggesting that Acb DA depletion has effects that cannot be attributed purely to motor deficits. These results are compatible with the loss of a dopaminergic motivational influence that contributes to normal performance. Indeed, Acb dopamine depletion does not only impair responding under instrumental reinforcement schedules, but also displacement behaviour occurring when food is delivered on a fixed-time schedule (Robbins & Koob, 1980). Such behaviour cannot easily be described as carrying a response cost, whereas it may reflect a potentiation of irrelevant available behaviours by a motivational effect of the food (Robbins & Koob, 1980).

The interpretation that the Acb contributes Pavlovian conditioned motivation to behaviour is compatible with the view that it mediates aspects of preparatory behaviour, temporally distant from the goal of behaviour (as opposed to consummatory behaviour, temporally close to the goal). As an example of such a distinction, lever-pressing by male rats for access to a female has been doubly dissociated from unconditioned sexual behaviour (Everitt *et al.*, 1987; Everitt & Stacey, 1987). This distinction has been phrased in various ways — preparatory and consummatory (Blackburn *et al.*, 1987; Robbins & Everitt, 1992), seeking and taking (Arroyo *et al.*, 1998; Everitt *et al.*, 1999), and sign tracking and goal tracking (Hearst & Jenkins, 1974). Manipulations of the Acb, including 6-OHDA lesions and systemic injections of dopamine receptor antagonists, have been shown to reduce the preparatory aspects (including rate of responding) of behaviour directed towards both food and (in male rats) a sexually receptive female, whilst leaving consummatory behaviour unaffected (Blundell *et al.*, 1977; Koob *et al.*, 1978; Kelley & Stinus, 1985; Blackburn *et al.*, 1987; Everitt, 1990). Schedule-induced polydipsia (SIP), a phenomenon whereby excessive drinking is produced by the intermittent presentation of small amounts of food, is also disrupted

selectively by 6-OHDA lesions of the Acb, but not of the dorsal striatum (Robbins & Koob, 1980; Mittleman *et al*., 1990). Acb lesions abolish SIP, leaving drinking / ingestion intact, whilst lesions of the dorsal striatum do not affect SIP but impair the ability of animals to drink effectively. In almost all paradigms studied, manipulations of limbic corticostriatal circuitry affect preparatory but not consummatory behaviour (Robbins & Everitt, 1992). The functional importance of Acb-dependent preparatory behaviour has also been demonstrated in a naturalistic setting by Whishaw & Kornelsen (1993). Rats normally carry food to a refuge to eat it, and when sated, carry the remaining food to hoard; rats with ibotenic acid or 6-OHDA lesions of the Acb were selectively impaired in this preparatory behaviour, failing to carry food to hoard it. The same rats were not impaired at carrying-to-eat, or eating itself.

Finally, a wide range of other tasks that depend on the effect of Pavlovian stimuli on instrumental or approach behaviour are also sensitive to lesions of the Acb or its afferents. The level of instrumental lever pressing is reduced by excitotoxic lesions of the Acb (Balleine & Killcross, 1994), consistent with the loss of a Pavlovian motivational effect that normally potentiates responding. Kelley *et al*. (1997) have also demonstrated a profound effect of intra-Acb infusions of glutamate receptor antagonists on Pavlovian and instrumental responding. Bilateral lesions of the BLA, or Acb, or a disconnection of the two, abolish a previously acquired conditioned place preference (CPP) for food (Everitt *et al.*, 1991); similarly, lesions of structures downstream from the Acb, including the ventral pallidum and mediodorsal thalamus, impair acquisition of a CPP (McAlonan *et al*., 1993). The BLA and Acb are also critical for the acquisition of responding under second-order schedules of sexual or cocaine reinforcement (Everitt *et al.*, 1989; Whitelaw *et al.*, 1996), in which the second-order CS is critical for responding in normal animals (Arroyo *et al.*, 1998) (for similar studies using heroin reinforcement, see Robbins *et al*., 2000; Alderson *et al*., in press-a; Alderson *et al*., in press-b). As discussed above (p. 42), lesions of the ACC, or AcbC, or a disconnection of the two, impair the acquisition of autoshaping (Bussey *et al.*, 1997a; Parkinson *et al.*, 2000c).

Summary

The nucleus accumbens is a key site mediating the ability of a Pavlovian conditioned stimuli to invigorate and direct behaviour; it is critical for autoshaping, the effect of psychostimulant-amplified conditioned reinforcers on instrumental responding, and PIT. This motivational influence of Pavlovian CSs has been termed *incentive salience* (Robinson & Berridge, 1993; Berridge & Robinson, 1998), or 'Pavlovian incentive value' (Dickinson *et al.*, 2000), to distinguish it from the instrumental incentive value of Dickinson and colleagues (Dickinson, 1994; Dickinson & Balleine, 1994) (*q.v.* and see pp. 24–27 for discussion of the differences between the two).

DELAYED REINFORCEMENT

Delayed reinforcement is of interest from two theoretical perspectives. Firstly, how do animals succeed in bridging delays to reinforcement at all? Natural reinforcers always follow the action that obtained them by a delay, even if it is short. Thus, to control the world successfully, animals must be able to use delayed reinforcement. In some species, the delay to reinforcement may be very long indeed; humans routinely make decisions on the basis of outcomes that are decades away. Secondly, what accounts for differences in individuals' ability to choose delayed rewards? Why are some individuals impulsive in their choices? These questions will be considered in order.

Delayed reinforcement in learning

Early theorists considered the fundamental problem of delayed reinforcement: how a response can be strengthened by reinforcement that follows it. Hull (1932) postulated that the strength of an S–R association is inversely related to the delay between the response and the reinforcement, assuming a logarithmic relationship. Indeed, instrumental learning has repeatedly been shown to be a decreasing function of the delay (e.g. Lattal & Gleeson, 1990; Dickinson *et al*., 1992). In several of the early studies, the delay was bridged by distinctive cues or environments. The cue that precedes eventual reward has the potential to become a secondary or conditioned reinforcer; thus the 'underlying' delay gradient function was unclear. In an effort to minimize the contribution of conditioned reinforcement, Grice (1948) trained rats on a visual discrimination task with delayed reinforcement. The rats had a choice of a white or a black start alley (which varied in their left / right position); the delay was provided by two grey alleys of variable length which terminated in two grey goal boxes. Choosing white led to a goal box with food; choosing black led to an empty box. Grice found that learning was noticeably impaired by as short a delay as 0.5 s, and severely impaired by 5 s. This deficit could be ameliorated by having more discriminable (black and white) goal boxes, or forcing the rats to make discriminable motor responses (climbing an incline or dodging between blocks) in the black and white start alleys.

Grice argued that Hull's primary delay of reinforcement did not exist and that learning under conditions of delayed reward was due to immediate secondary reinforcement, based on traces of visual or proprioceptive stimuli. Clearly, if the primary gradient does exist, it is steep; the distinction becomes one of whether the delay applies to response reinforcement (Hull) or stimulus–reward association (Grice).

One other perspective deserves comment: that of Killeen & Fetterman (1988), who suggested that the very idea of a 'delay gradient' is misleading. In their model, reinforcement always strengthens the responses that the animal is presently making, and never acts 'backwards in time' to strengthen past responses. The observed 'gradient' stems from the fact that the animal has a finite probability of leaving the behavioural state it was in when it responded; if reinforcement follows immediately, there is a high probability of strengthening the response that caused reinforcement, but the longer the reinforcer is delayed, the greater the chance that the animal has moved to another state, in which case a different response will be reinforced. This point has also been made by Spence (1956), Mowrer (1960), and Revusky & Garcia (1970); see also Mackintosh (1974, pp. 155–159).

It is obviously impossible for response–reinforcement or stimulus–reinforcement learning to occur unless the trace of the response or the stimulus persists to be reinforced or associated. Whichever of the three perspectives has most merit, the point is made that small delays of reinforcement can markedly impair learning, that stimuli differentially associated with reward can improve this performance, and that interoceptive cues can sometimes perform this function.

Choice, and pathological choice, from the perspective of utility theory

Before considering the role of delayed reinforcement in choice behaviour, I will briefly review one theoretical approach to choice behaviour, utility theory, that has explicitly or implicitly underlain many studies using delayed reinforcement.

Utility theory

Formal utility theory is based on six axioms that define attributes of preference that perfectly rational agents should possess (von Neumann & Morgenstern, 1947) (reviewed by Russell & Norvig, 1995). One, for example, is *transitivity*: if an agent prefers *A* to *B* and *B* to *C*, then it must prefer *A* to *C*. If the agent

violated this principle, preferring $A>B>C>A$, then an observer could offer the agent C in exchange for A and a small monetary payment; similarly B for C and A for B, after which the agent ends up in its original state but with less money, which (assuming money is desirable) is irrational.

Given that a rational agent obeys these axioms, then there must exist a *utility function* U that assigns a real number to every outcome O such that $U(O_1) > U(O_2)$ if O_1 is preferred to O_2, and $U(O_1) = U(O_2)$ if the agent is indifferent between the two outcomes.

Goal-directed action requires that the agent assigns value (goal status) to outcome states, but also that it knows the consequences of its actions. To allow for the fact that actions may not always have totally predictable consequences, the agent's knowledge about the causal nature of the world may be represented in the form $p(action \rightarrow outcome_n \mid evidence)$ denoting the probability, given the available evidence, that *action* causes $outcome_n$. The *expected utility* of an action is therefore given by:

$$EU(action \mid evidence) = \sum_n p(action \rightarrow outcome_n \mid evidence) \cdot U(outcome_n)$$

Rational decision-making follows if the agent selects the action with the maximum expected utility (the MEU principle). The theory specifies neither the utility functions themselves — anything can be valued — nor the way that the decision is arrived at, which may be explicit or implicit.

The formal decision-making approach described suffers from two particular deficiencies. Firstly, computing the expected utilities takes finite time. It may often be better to make an imperfect decision quickly than *eventually* to make what *would* have been the perfect decision. In artificial intelligence, this has proved a difficult problem (Russell & Norvig, 1995). Secondly, the MEU principle implies that in identical situations, the same action will always be taken (it is a 'pure' strategy). However, game theory (von Neumann & Morgenstern, 1947) has shown that there are many situations involving choice under uncertainty when the optimal strategy is to assign probabilities to making different choices but to let the actual decision be governed by chance (a 'mixed' strategy). Even using this method, Gödel's (1931) incompleteness theorem implies that no concept of rationality can be optimal in every situation. How randomness is used in decision-making is poorly understood; Mérö (1998) provides an entertaining look at these issues.

Pathological choice in the context of utility theory

There are two ways within the framework of utility theory to produce 'pathological' decision-making. One is to alter the utility functions. For example, assigning a higher utility to poverty than wealth would cause a perfectly rational agent to give its money away; if gambling had intrinsic utility then an agent might gamble despite the financial loss. While the underlying choice remains rational, the agent's preferences generate abnormal behaviour. Indeed, some investigators see it as axiomatic that animals make rational or optimal decisions (see Williams, 1994b, pp. 91/94), so that the experimenter's job is to discover the value system of the subject.

The other mechanism, considered less often, is that utilities are computed normally but the decision-making process itself fails. Indeed, normal humans are not 'normative': they systematically deviate from the axioms of decision theory (Kahneman *et al*., 1982; see also Chase *et al*., 1998), which, incidentally, is why computerized systems can outperform human experts (Heckerman *et al*., 1992).

The distinction is difficult. As an illustration, consider a smoker who desires to abstain but lights a cigarette. Are we to consider the decision flawed or the actual utility of smoking higher than he thought? If 'optimality can be considered axiomatic' (Williams, 1994b, p. 94), the latter is the case, but such a theory cannot distinguish between the act of our relapsing smoker and one who has no wish to give up. Nev-

ertheless, the distinction seems important; these questions only begin to make sense within a reductionist approach to the way the brain reaches decisions.

Views of choice in the brain

To choose between two goals that differ in nature, such as food *v.* money, they must be compared on a single dimension. Utility functions achieve this by converting multifactorial alternatives to real numbers. Neurally, a similar process is logically unavoidable — if at no earlier stage of processing, incompatible behaviours must compete for access to motor output structures (although there is no *a priori* reason why the neural comparison process should be simple or linear).

There is a long history of behavioural research into the computation of reward utility and consequent behavioural strategy (reviewed by Williams, 1994b), including the utility of artificial reinforcers (see Shizgal, 1997). One approach used was to calculate the efficacy of reinforcement by establishing the relationship between response *rate* and the frequency and amount of reinforcement; however, such attempts soon established that this relationship was not simple (see Williams, 1994b, pp. 82–83). For example, response rates are affected by whether a ratio or an interval schedule of reinforcement is used, even when the reinforcement rate is identical (Dawson & Dickinson, 1990). Similarly, the mechanisms governing motor aspects of responding are neurally dissociable from motivational mechanisms (see e.g. Robbins & Everitt, 1992).

The matching law and related research: 'top-down' analyses of choice

Another approach has been to relate reinforcement efficacy to choice behaviour. This literature stems from the discovery by Herrnstein (1961; 1970) of the 'matching law'. Herrnstein (1961) trained pigeons to respond on two concurrent variable interval (VI) schedules, and varied the relative availability of reinforcement on the two schedules while holding the overall reinforcement rate constant. He observed that the proportion of the total behaviour allocated to each response key approximately matched the proportion of reinforcers allocated to that key. This defines the matching law:

$$\frac{R_1}{R_1 + R_2} = \frac{r_1}{r_1 + r_2}$$

where R represents the behavioural response rate for each alternative, and r the reinforcement. Herrnstein (1970) extended this relationship to take account of more than two alternatives, particularly including 'unmeasured' activities the animal may engage in, and derived a 'general principle of response output' (Herrnstein, 1970, p. 256):

$$R_1 = \frac{kr_1}{r_1 + r_e}$$

where R_1 is the rate of the response being measured, r_1 is the quantity of reinforcement for that response, r_e is the reinforcement for all other responses, and k is a parameter determining the maximum response rate. Although there are situations where the matching law is not useful — in particular, ratio schedules, where the distribution of reinforcement necessarily *follows* the distribution of responding — a vast literature has sought to define the effects of varying parameters of reinforcement (such as rate, probability, delay, and magnitude) based on this work (see de Villiers & Herrnstein, 1976).

Problems have emerged. In many circumstances, subjects have been found to 'overmatch' (exhibit preferences that are exaggerated relative to the predictions of the matching law) or 'undermatch' (exhibit reduced preferences), requiring further development of the mathematical models (Baum, 1974; Baum, 1979), though it has been argued that this is a circular approach (Rachlin, 1971). Maximum response rates (k in the equation above) have been shown to vary with the kind of reinforcement used (Belke, 1998),

violating an assumption of Herrnstein's law. Nevertheless, the matching law and its extensions do a good job of describing the relationship between reinforcement rate and behaviour on concurrent VI and concurrent-chain schedules (Williams, 1994b).

The matching law described a molar property of behaviour — that is, the overall distribution of a large number of responses. As responses are made on a moment-to-moment basis, the question arises of what 'molecular' choice process operates to produce matching at a molar level. Suggestions vary from 'momentary maximizing' theory, which suggests that subjects choose (in all-or-none fashion) the response with the highest instantaneous reinforcement probability, to the idea that matching is the basic choice rule (see Mackintosh, 1974, pp. 192–195; Williams, 1994b).

Relating choice to 'value'

All these theories share a theoretical basis: it is assumed that some value is computed for each alternative behaviour, and a single decision rule allocates behaviour according to the relative distribution of values. In order to produce a single value for each alternative, different reinforcement parameters (rate, magnitude, delay, and probability) converge on a single dimension (Baum & Rachlin, 1969). Often, the effects of these different parameters are assumed to be calculated independently (Killeen, 1972; Rachlin *et al*., 1991; Ho *et al*., 1999). Though some investigators have supported the latter assumption (Mazur, 1987; Mazur, 1997), using different techniques, others have found that the effects of reinforcer delay and magnitude are not independent (Ito, 1985; White & Pipe, 1987). In either case, the assumption that all choice alternatives are reduced to a single value and then compared in order to select the option with the greatest value corresponds directly to a form of utility theory, as described above.

Fragmenting choice: a neuropsychological or 'bottom-up' approach

We have seen how utility theory can fail to characterize human decision-making (Kahneman *et al.*, 1982), just as similar approaches have not fully characterized choice in other animals (Williams, 1994b, p. 105). Perhaps more success can be achieved by considering the multiple psychological systems that have been discovered to contribute to instrumental performance. In this framework, behaviour and choice are seen as the asymptotic sum of contributions from cognitive goal-directed systems, habitual responding and other motivational influences (e.g. Dickinson, 1994). As we have seen, rats possess *at least* two representations of the value of foodstuffs (Dickinson & Balleine, 1994), namely hedonic value and the incentive value governing instrumental responding; Pavlovian incentive value is probably a third (see pp. 24 & 28). An analysis of the neuropsychological mechanisms by which these multiple motivational systems calculate the value of environmental events and interact with each other may prove more productive than the 'top-down' approach. To take a hypothetical example, suppose that stimulus–response habits obey the matching law, but that cognitive, voluntary decisions can override habits in some circumstances and have a different value system. It is likely that acknowledging the existence of these two systems, and determining when each operates, will more rapidly lead to an accurate description of choice behaviour than attempting to model choice with a single, but highly complicated, value system.

Neuropsychological research along these lines is a young field. As has been outlined above, consideration of the neural basis of Pavlovian and instrumental conditioning in animals has led to the identification of several brain regions and neurotransmitter systems that are involved in reinforcement and value assessment (including the reinforcing effects of drugs of abuse). This literature supports neuropsychological data derived from studies of humans with acquired disorders of decision-making. Noteworthy among these are studies of humans with damage to the orbitofrontal cortex or amygdala (including the famous case of orbitofrontal cortex damage in Phineas Gage, first reported by Harlow, 1868), who exhibit im-

paired choice behaviour despite apparent knowledge that they are choosing poorly (Damasio, 1994; Bechara *et al*., 1998). Following the pioneering theories of Nauta (1971), this work has led to the development of a specific theory of a process contributing to choice, namely the somatic marker hypothesis (Damasio, 1994; Damasio, 1996; Bechara *et al*., 2000). This theory proposes the existence of a non-conscious, rapidly-retrieved utility signal that improves decision-making performance by removing poor options from the consideration of a computationally-intensive cognitive process. These signals appear to have a measurable autonomic correlate in galvanic skin responses (Bechara *et al*., 1996; Bechara *et al*., 1997) and depend upon the integrity of the amygdala and orbitofrontal cortex (Bechara *et al*., 1998; Bechara *et al*., 1999). This may represent the best-characterized neural correlate of decision-making.

The other major avenue of investigation into pathological decision-making has concentrated on the phenomenon of impulsivity. Research into impulsivity is deeply interwoven with the study of delayed reinforcement; these areas will be reviewed next.

Impulsivity and impulsive choice

Impulsivity was well known to the ancient Greeks. The character flaw *akrasia* (weakness of will, lack of self-control, or incontinence) is a deficiency of the power to act as one judges best in the face of competing motivation. Aristotle saw it as a commonplace deviation from the norm of men:

> *'The incontinent man, knowing that what he does is bad, does it as a result of passion, while the continent man, knowing that his appetites are bad, refuses on account of his rational principle to follow them.'*
>
> Nicomachean Ethics (Aristotle, 350 BC / 1925, book 7, chapter 1)
>
> *'It is plain, then, that incontinent people must be said to be in a similar condition to men asleep, mad, or drunk.'*
>
> (Book 7, chapter 3.)
>
> *'Now incontinence and continence are concerned with that which is in excess of the state characteristic of most men; for the continent man abides by his resolutions more and the incontinent man less than most men can.'*
>
> (Book 7, chapter 10.)

Aristotle's definition of incontinence focuses on an inability to suppress one's desires in favour of more rational, high-minded resolutions. However, the term 'impulsivity' has been applied to many different aspects of maladaptive choice. Ainslie (1975) summarized three guesses about why humans are prone to obey 'impulses':

(1) that they lack insight into the consequences of their actions — a defect in instrumental contingency learning;
(2) they are aware of the consequences of their actions, but are unable to suppress 'some lower principle (the devil, repetition compulsion, classical conditioning)' — a defect in response inhibition;
(3) they are aware of the consequences of their actions, and choose rationally according to their value system, but their values are 'innately distorted so that imminent consequences have a greater weight than remote ones' — reduced value of delayed reinforcement.

Impulsivity may be given a broader scope still; Evenden's (1999b) review of the field encompasses all the above and adds 'preparation impulsivity' (reaching a decision before adequate information is gathered) and 'execution impulsivity' (interrupting a chain of behaviour before its goal is achieved) (Evenden, 1998, p. 37). Critically, these aspects of impulsivity may be dissociated pharmacologically, implying that they reflect genuinely different underlying processes (Evenden, 1999b).

Impulsivity may be considered a normal personality trait (Eysenck & Eysenck, 1977; Barratt & Patton, 1983; Eysenck, 1993), as Aristotle did, but it is also a feature of a number of clinical disorders. These include personality disorders (antisocial personality disorder and borderline personality disorder; APA, 1994), impulse control disorders, including drug addiction ('substance abuse disorder'; APA, 1994), and attention-deficit / hyperactivity disorder (ADHD), a prevalent disease of childhood of which impulsivity is one sign (Evenden, 1998; Sagvolden & Sergeant, 1998).

Multifaceted though impulsivity is, the present thesis will focus exclusively on *impulsive choice*, exemplified by the inability of an individual to choose a large delayed reward in preference to a small immediate reward (or, in an aversive context, an inability to choose a small immediate penalty in preference to a large delayed penalty). This form of impulsivity can be characterized as pathological hypersensitivity to delays of reinforcement (though the formally identical aversive analogy may make it clearer that impulsive choice may also have something to do with relative insensitivity to differences in reinforcer magnitude). What, though, is 'normal' sensitivity to delayed reward?

Delayed reinforcement in choice

In a typical situation, a subject chooses between an immediate, small reward or a large, delayed reward; the time discounting function quantifies the effect of the delay on preference. Kacelnik (1997) points out that economic models of choice tend to be based on exponential time discounting functions. If the starting assumption is that delayed reward is preferred less because there is a constant probability of losing the reward per unit of waiting time, or that there is a constant 'interest rate' for the reward obtained immediately (and that the subject's behaviour is attuned to this fact, i.e. that choice is normative) then exponential models emerge. If a delayed reward of magnitude A is chosen and there is a probability p of loss in every unit of time waited, the perceived value V of the delayed reward should be $V = A(1 - p)^T = Ae^{-kT}$ where $k = -\ln(1 - p)$.

However, the exponential model has been emphatically rejected by experimental work with humans and other animals. The literature on human cognitive decisions will not be considered here. The rat literature contains several demonstrations (many based on the adjusting-delay task of Mazur, 1987) procedure, using natural reinforcers and intracranial self-stimulation (or 'brain-stimulation reward') (Grice, 1948; Mazur, 1987; Mazur *et al*., 1987; Richards *et al*., 1997b), that time discounting is described well by a *hyperbolic* discount function (Figure 9) or at least a very similar power law (Grace, 1996). Kacelnik (1997) offers some explanations as to why hyperbolic discounting may be in some sense optimal. One interesting prediction from this function is that preference between a large and a small reward should be observed to reverse depending on the time that the choice is made (Figure 10), and such preference reversal is a reliable experimental finding (for references see Bradshaw & Szabadi, 1992). Of course, the neuropsychological system responsible for hyperbolic discounting is unknown — such discounting might, for example, result from poor knowledge of the action–outcome contingency at long delays, from weak stimulus–response habits, or from reduced utility of delayed rewards in the context of perfect contingency knowledge.

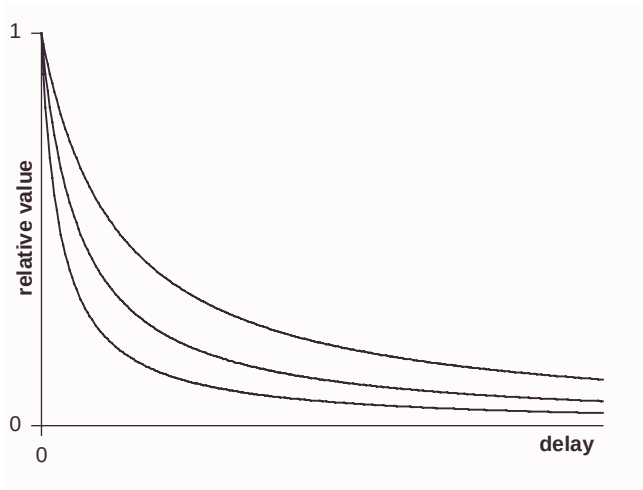

Figure 9. Hyperbolic discounting, governed by the equation

$$V = \frac{magnitude}{1 + K \cdot delay}$$

Large values of K give the steepest curve.

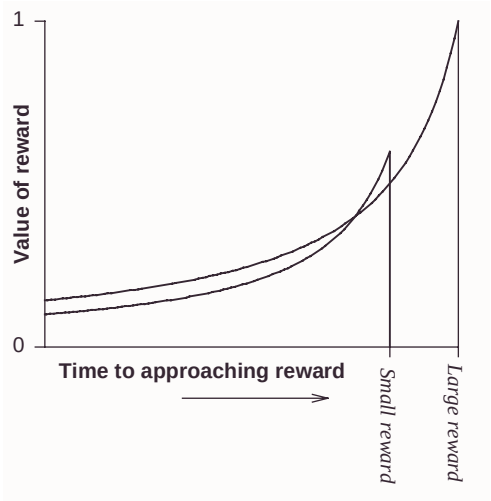

Figure 10. Preference reversal. If given a choice between an early reward of value 0.6 and a later reward of value 1, hyperbolic discounting predicts that the larger reward will be chosen if the choice is made far in advance (towards the left of the graph). However, as time advances, there comes a time just before delivery of the small reward when preference reverses and the small reward is chosen. Figure adapted from Ainslie (1975).

Neurochemical and neuroanatomical studies of delayed reinforcement

Serotonin (5-HT)

Abnormalities of the utility function for delayed reinforcement have been suggested to occur following neurochemical manipulations. The suggestion that serotonin is involved in impulse control follows from the twin observations that drugs that suppress 5-HT function appear to reduce behavioural inhibition, making animals more impulsive in the 'motor' sense (Soubrié, 1986), and that low levels of serotonin metabolites in cerebrospinal fluid are associated with impulsive aggression and violence in humans (e.g. Åsberg *et al*., 1976; Linnoila *et al*., 1983; Brown & Linnoila, 1990; Linnoila *et al*., 1993) and risk-taking behaviour in monkeys (Mehlman *et al*., 1994; see also Evenden, 1998). In the sphere of delayed reinforcement, forebrain serotonin depletion, which leads to 'impulsive choice' in a variety of paradigms (Wogar *et al*., 1993b; Richards & Seiden, 1995; Bizot *et al*., 1999), has been suggested to reflect a modification of the temporal discounting function (Wogar *et al.*, 1993b; Ho *et al.*, 1999). Specifically, 5-HT depletion is suggested to steepen the function, such that delayed rewards lose their capacity to motivate or reinforce behaviour. The animal becomes hypersensitive to delays (or hyposensitive to delayed reward). As delayed rewards have unusually low utility, the animal consistently chooses small, immediate rewards over large, delayed rewards, a characteristic of impulsivity (Ainslie, 1975). The specific contribution of different 5-HT receptor subtypes to choice of delayed reward has also been studied (Evenden & Ryan,

1996; Evenden, 1998; Bizot *et al.*, 1999; Evenden, 1999b; Evenden & Ryan, 1999), but this topic will not be pursued in detail.

Dopamine and attention-deficit/hyperactivity disorder (ADHD)

Much of the interest in the relationship between dopamine and impulsivity comes from the discovery that amphetamine and similar psychostimulants are an effective therapy for ADHD (Bradley, 1937). Though these drugs have many actions, they are powerful releasers of dopamine from storage vesicles in the terminals of dopaminergic neurons, and prevent dopamine re-uptake from the synaptic cleft, potentiating its action (for references see Feldman *et al*., 1997, pp. 293/552/558). Sagvolden & Sergeant have proposed that many features of ADHD, including preference for immediate reinforcement and hyperactivity on simple reinforcement schedules (due to short inter-response times; Sagvolden *et al*., 1998), are due to an abnormally short and steep delay gradient and that this is due to a hypofunctional dopamine system. Indeed, they go on to suggest Acb DA as the specific culprit (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998). Clearly, accumbens dopamine is implicated in aspects of responding for reinforcement, as discussed earlier, though its role is not yet fully understood.

Many of the inferences regarding the neural abnormalities in children with ADHD have in fact been drawn from studies of the spontaneously hypertensive rat (SHR), an inbred strain of rat that serves as an animal model of ADHD (Wultz *et al*., 1990; Sagvolden *et al*., 1992; Sagvolden *et al*., 1993; Sagvolden, 2000). This rat exhibits pervasive hyperactivity and attention problems that resemble ADHD, is abnormally sensitive to immediate reinforcement in the sense that it exhibits a steeper 'scallop' of responding on fixed-interval (FI) schedules (Sagvolden *et al.*, 1992), and is impulsive on measures of 'execution impulsivity' (Evenden & Meyerson, 1999).

Examination of the brains of SHRs supports the assertion that they have an abnormality of dopamine systems. Depolarization- and psychostimulant-induced dopamine release in nucleus accumbens brain slices is altered in the SHR compared to Wistar Kyoto (WKY) progenitor control rats in a complex pattern that has been attributed to hypofunction of the mesolimbic dopamine system (de Villiers *et al*., 1995; Russell *et al*., 1998; Russell, 2000), though abnormalities have also been found in dopamine release in slices of dorsal striatum and prefrontal cortex (Russell *et al*., 1995). Amygdala dysfunction has also been suggested (Papa *et al*., 2000). Within the Acb, differences in gene expression and dopamine receptor density have been observed in both the core and shell subregions (Papa *et al*., 1996; Carey *et al*., 1998; Papa *et al*., 1998).

Systemic psychopharmacological studies using normal animals provide an additional source of evidence regarding dopamine systems and choice of delayed reward. Many of these studies have examined the role of psychostimulant drugs such as amphetamine and methylphenidate, given these drugs' efficacy in treating some symptoms of ADHD. However, conflicting results have been obtained in animal models. For example, psychostimulants have sometimes been found to promote choice of delayed rewards, and sometimes to impair it, in normal rats (Sagvolden *et al.*, 1992; Charrier & Thiébot, 1996; Evenden & Ryan, 1996; Richards *et al*., 1997a; Richards *et al*., 1999; Wade *et al*., 2000). These apparent inconsistencies will be addressed in Chapter 6 by considering the potential contribution of conditioned reinforcement to the effects of psychostimulants in altering preference for delayed reward.

The prospect of delineating neural circuitry involved in choice of delayed reward

Although there has been considerable research on the neurochemical basis of tolerance to delayed rewards, together with correlative studies of cortical functional abnormalities in ADHD children and of regional differences in neurotransmission in the SHR, there have been few direct investigations of the role

of anatomically-defined brain structures in the capacity of animals to choose a delayed reward. Dopamine and serotonin affect this capacity, but where do they have their action? Abnormal functioning of prefrontal cortical regions, including the ACC, has been observed in humans with ADHD (Ernst *et al*., 1998; Bush *et al*., 1999), but it is not known whether these abnormalities are related to impulsive choice. In Chapter 7, I will consider the role of the ACC, the mPFC, and the AcbC, three structures that play a role in reinforcement processes, as outlined above, and that also receive serotonergic and dopaminergic projections (see Fallon & Loughlin, 1995; Halliday *et al*., 1995).

Tasks used to study choice of delayed reinforcement

Two main approaches have been used to study delayed reinforcement: free-operant and discrete-trial procedures.

Free-operant tasks are typified by the concurrent-chain schedule (Autor, 1969; see Davison, 1987). In the most common variant, subjects respond on two concurrent VI schedules (the 'initial links'). When one of the two schedules is completed, the other manipulandum is deactivated and the chosen schedule enters the 'terminal link' (in which reinforcement is provided on another schedule — for example, a fixed time schedule, in which noncontingent reinforcement is given after a fixed delay). Relative response allocation in the initial links is taken as a measure of relative preference for the two terminal link schedules. Versions of concurrent chain schedules that depend on subjects' timing behaviour have also been developed (Gibbon & Church, 1981).

Though they allow accurate determination of relative response rates and, by inference, relative preference, such free-operant schedules carry two problems of interpretation. One is that the delays between initial-link responses and initiation of the terminal link may also form part of the delay to reinforcement; such delays are difficult to control for, and their importance may vary with the relative durations of initial and terminal links. The other applies to pharmacological and lesion studies: manipulations that affect an animal's ability to produce motor responses, to switch between two responses, and to time their motor output, may all confound interpretation of the results (Ho *et al.*, 1999). For these reasons, free-operant schedules were not used to assess preference for delayed reward in the present thesis.

Discrete-trial schedules may also be divided into two classes. In the simplest type, the subject chooses between two mutually exclusive alternatives and preference is measured as the proportion of trials on which each alternative is chosen (e.g. Bradshaw & Szabadi, 1992; Evenden & Ryan, 1996). Choice in a T-maze has been used similarly (e.g. Bizot *et al*., 1999). Although such schedules may not provide as accurate a measure of preference, as subjects tend towards exclusive preference on discrete-trial and ratio schedules (see Mackintosh, 1974, pp. 190–195), the response–reinforcer delay can be accurately controlled and instantaneous choice is free of the confounds discussed above regarding response rates and switching rates.

An alternative type of discrete-trial task, the adjusting-delay schedule, was invented by Mazur (1987). In this task, subjects choose in discrete trials between a *fixed* alternative, such as a small immediate reward, and an *adjusting* alternative, such as a larger reward delivered after a delay. This delay can alter. If the subject prefers the larger reward, the delay is lengthened, and if it prefers the smaller alternative, the delay is shortened, in an attempt to titrate the subject's preference towards indifference. At this indifference point, the length of the adjusting delay is taken as a measure of the subject's preference between the two reinforcers — it is the delay that 'balances' the difference in the magnitudes of the two reinforcers. (A similar schedule in which the amount of reinforcement is adjusted was recently described by Richards *et al.*, 1997b). The key advantage of the indifference-point methodology is that it allows quantitative es-

timation of subjects' preferences without assuming a particular relationship between reinforcer value and behavioural output — the only assumption required is that when two reinforcers are of equal value, behaviour is equally distributed between the two response alternatives. The technique also allows subjects' sensitivity to reinforcer delay to be distinguished from sensitivity to reinforcer magnitude (see Ho *et al.*, 1999). However, the schedule design is complex: not only does the delay affect the subject's choice, but choice affects the adjusting delay. Versions of this schedule have been used to assess the effects of chronic neurochemical manipulations (Wogar *et al.*, 1992; Wogar *et al.*, 1993b; Ho *et al.*, 1997). In Chapter 5, a group of rats are tested on a version of the adjusting-delay schedule to see if the task is suitable for other neurotoxic lesion and acute pharmacological studies.

It is worth noting at this point that there is a great difference between measuring delay preference using tasks in which trials occur at a fixed frequency (or at least, in which the subject's choice does not influence the time to the next choice-point) and those where choice can influence this frequency. In the latter kind of task, the subject may choose the small immediate reinforcer and have the opportunity to do so again very rapidly, so may be able to accumulate more reward by repetitive choice of the small reinforcer than by choosing the large one. If the overall trial frequency is held constant, however, the strategy that maximizes reward is always to choose the larger reinforcer; in this case, failure to do so can be attributed to its delay. Some studies of impulsive choice in children with ADHD used the former type of delay-of-gratification task, finding impairments, while comparable studies with a fixed trial length have failed to find differences (see Sonuga-Barke *et al.*, 1998). All studies reported in this thesis used a fixed trial frequency.

Conditioned reinforcement in choice of delayed rewards

Finally, when considering preference for delayed reinforcement, the role of conditioned reinforcers must be considered. Not infrequently, delay-of-reinforcement procedures have been used as a tool to study conditioned reinforcement; a stimulus is presented during the delay to reinforcement in the expectation that it will become a conditioned reinforcer. Such stimuli certainly affect choice behaviour (see Lattal, 1987), and tasks of this sort have served as the basis for attempts to quantify the value of conditioned reinforcers (Autor, 1969; Mazur, 1991; Mazur, 1995; Mazur, 1997), though the issues are complex (Williams, 1994a).

For the present thesis, it suffices to note that conditioned reinforcement can be an important factor influencing preference for delayed reward. In Chapters 6 and 7, I will seek to clarify the effects of certain drugs, including amphetamine, that are known to affect responding for conditioned reinforcement. In these experiments, explicit comparison will be made between the situation in which a stimulus is present during the delay to reinforcement and the situation in which no such stimulus is present. In Chapter 7, when lesion studies are conducted, the tasks used will not present a stimulus during this delay, to avoid this potential problem of interpretation.

ORGANIZATION OF EXPERIMENTAL WORK IN THIS THESIS

The experiments described in this thesis may be divided into two parts.

In Part 1 (Chapters 3 & 4), a clearer understanding is sought of the role of the ACC and Acb in basic Pavlovian and instrumental processes. In Chapter 3, I will consider in detail the functions of the ACC, as elucidated by previous rodent and primate studies, before investigating its contribution to simple Pavlovian conditioned approach, conditioned reinforcement, Pavlovian–instrumental transfer, and other simple

conditioning procedures. In Chapter 4, the involvement of the core and shell subdivisions of the Acb in Pavlovian–instrumental transfer is investigated, using a task with greater behavioural specificity than those previously used for this purpose.

In Part 2 (Chapters 5–7), an attempt is made to understand the contributions of limbic corticostriatal circuitry to the capacity of rats to choose delayed reward. In Chapter 5, a detailed examination is made of rats' performance on the adjusting-delay schedule of Mazur (1987), and this schedule is found unsuitable for pharmacological or lesion studies. In Chapter 6, a different task is developed, from that of Evenden & Ryan (1996). This task is used to investigate the effects of systemic dopaminergic drugs on choice of delayed reinforcement, and their interactions with stimuli that 'bridge' the delay to reinforcement. In Chapter 7, the same task is used to investigate the effects of destroying key elements of the limbic corticostriatal circuit — the ACC, mPFC, and AcbC — on choice of delayed reward, and the effects of intra-accumbens injections of amphetamine.