Choosing delayed rewards: perspectives from learning theory, neurochemistry, and neuroanatomy

Rudolf N. Cardinal, Trevor W. Robbins & Barry J. Everitt

Department of Experimental Psychology, University of Cambridge,
Downing Street, Cambridge CB2 3EB, UK.
E-mail: Rudolf.Cardinal@pobox.com
Telephone: +44–7092–340641. Fax: +44–7092–340645.

Submitted for *Choice, Behavioral Economics, and Addiction*
(Birmingham, Alabama, USA; 15–17 March 2002)

Revised version: 13 April 2002

1 . INTRODUCTION 1

2 . THE RAT RESPONDS FOR REWARDS BASED ON MULTIPLE REPRESENTATIONS OF REINFORCER VALUE 2

3 . MECHANISMS BY WHICH DELAYED REINFORCEMENT CAN AFFECT INSTRUMENTAL LEARNING 4

4 . DELAYED REINFORCEMENT IN CHOICE 6

4.1 . UTILITY THEORY AS AN APPROACH TO NORMAL AND PATHOLOGICAL CHOICE IN ANIMALS 6
4.2 . CHOICE WITHIN THE BRAIN: 'TOP-DOWN' AND 'BOTTOM-UP' APPROACHES 8
4.3 . TEMPORAL DISCOUNTING AND IMPULSIVE CHOICE 11

5 . NEUROCHEMICAL STUDIES OF IMPULSIVE CHOICE 13

5.1 . SEROTONIN (5-HT) 13
5.2 . DOPAMINE (DA) 13

6 . NEUROANATOMICAL STUDIES OF IMPULSIVE CHOICE 17

6.1 . ROLE OF THE NUCLEUS ACCUMBENS (ACB) 18
6.2 . ROLE OF THE ANTERIOR CINGULATE CORTEX (ACC) 22
6.3 . ROLE OF THE MEDIAL PREFRONTAL CORTEX (MPFC) 23
6.4 . ROLE OF THE ORBITOFRONTAL CORTEX (OFC) 24

7 . CONCLUSIONS 24

7.1 . THEORIES OF LEARNING AND CHOICE WITH DELAYED REWARD 24
7.2 . NEURAL BASIS OF CHOOSING DELAYED REINFORCERS 28
7.3 . APPLICATION TO ADDICTION AND OTHER DISORDERS OF IMPULSIVITY 29

ACKNOWLEDGEMENTS 32

FIGURE LEGENDS 32

REFERENCES 35

1. Introduction

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, individuals differ in their ability to choose delayed rewards. Impulsive choice is exemplified by the tendency of an individual to choose a reward that is small or poor, but is available immediately, in preference to a larger reward that is only obtainable after a period of time (Ainslie 1975). Impulsive choice may reflect reduced efficacy of delayed reinforcement. It has been considered a normal human characteristic (Aristotle 350 BC / 1925), but impulsive choice contributes to deleterious states such as drug addiction (Poulos *et al.* 1995; Heyman 1996; Bickel *et al.* 1999; Evenden 1999a; Mitchell 1999) and has been suggested to underlie a number of other clinical disorders, including attention-deficit / hyperactivity disorder (ADHD) (Sagvolden *et al.* 1998; Sagvolden and Sergeant 1998). Why are some individuals impulsive in their choices?

In this chapter, as a background to these questions, modern theories of animal learning (specifically, Pavlovian and instrumental conditioning) will first be summarized. These theories are based on studies showing that multiple dissociable psychological processes contribute to rats' actions. The potential ways in which delayed reinforcement can affect instrumental *learning* will be considered. Theories of instrumental *choice* will then be briefly reviewed, examining their applicability to choice between delayed reinforcers and their relevance to neuropsychological studies. Interventional studies will be reviewed that examine the role of selected neurochemical systems (the serotonin and dopamine neuromodulator systems) and neuroanatomical regions (the nucleus accumbens core,

anterior cingulate cortex, medial prefrontal cortex, and orbitofrontal cortex) in rats' ability to choose delayed rewards. Finally, the applications of these studies to addiction and other disorders of impulsivity will be considered.

2. The rat responds for rewards based on multiple representations of reinforcer value

In addition to innate, unlearned behaviour, rats may acquire new behavioural responses through Pavlovian conditioning, in which the experimenter arranges a contingency between stimuli in the world by presenting those stimuli independently of the animal's behaviour (Pavlov 1927). Similarly, rats exhibit goal-directed behaviour learned through instrumental conditioning, in which the experimenter arranges a contingency between the animal's behaviour and a reinforcing outcome (Thorndike 1911). When considering subjects choosing between delayed rewards, it is natural to think of instrumental conditioning as the underlying psychological process. However, instrumental conditioning is not a unitary phenomenon. Apparantly simple goal-directed responding depends on multiple representations within the brain (Dickinson 1994; Dickinson and Balleine 1994). Additionally, under certain circumstances, such behaviour can cease to be goal-directed and become habitual (Adams 1982; Dickinson 1994), and Pavlovian conditioning can have a direct influence on instrumental responding. Indeed, it is apparent that many psychological processes contribute to learning and performance of instrumental behaviour (summarized here but reviewed thoroughly by Dickinson 1994; Dickinson and Balleine 1994; Cardinal *et al.* in press); thus, complex tasks such as choosing between delayed rewards may depend on a number of these processes.

Specifically, rats can learn the *instrumental contingency* between an action such as lever-pressing and an outcome (such as food). This knowledge is represented in declarative fashion in the brain (Dickinson 1994). To produce goal-directed action, this knowledge interacts with another declarative representation, that of the *instrumental incentive value* of

the food, in order to determine whether the rat does or does not perform the action. If the rat knows that lever-pressing produces food and values the food highly, it is likely to press the lever. Surprisingly, there is a second system that assigns a value to foodstuffs — the food's *affective* or *hedonic value*. Through direct experience, the hedonic value is 'written into' the instrumental incentive value governing goal-directed action (a process termed incentive learning: Dickinson 1994; Dickinson and Balleine 1994). There are circumstances in which these two values can differ (see Dickinson 1994; Dickinson and Balleine 1994; Cardinal *et al.* in press).

Although rats possess declarative knowledge of the consequences of their actions, they also possess a procedural, stimulus–response 'habit' system; this system is less flexible (for example, it contains no explicit knowledge of the reinforcer, so the probability of performing the habitual action does not change immediately if the value of the reinforcer is altered). Overtraining is a typical way in which a goal-directed response becomes habitual (see Adams 1982; Dickinson *et al.* 1983; Dickinson 1985; Dickinson 1994; Dickinson *et al.* 1995).

Finally, Pavlovian conditioned stimuli (CSs) can modulate instrumental performance directly (Dickinson 1994; Dickinson and Balleine 1994), an effect termed Pavlovian–instrumental transfer (PIT). For example, a Pavlovian CS that predicts the arrival of sucrose solution will enhance instrumental lever-pressing for sucrose (Estes 1948; Lovibond 1983); this, at least in part, represents *conditioned motivation* (see Cardinal *et al.* in press). The ability of a Pavlovian CS to affect instrumental performance depends upon the relevance of the unconditioned stimulus (US) to the animal's motivational state at the time; therefore, a neural system must exist to judge the current value (or salience) of the US when the CS is presented. This value has been dissociated both from the instrumental incentive value and the hedonic value of foodstuffs (Dickinson 1986; Dickinson and Dawson 1987a; Dickinson and Dawson 1987b; Dickinson *et al.* 2000; see Cardinal *et al.* in

press). As emphasized by Gjelsvik (this volume), cue-induced motivation (PIT) may have great significance for addiction (with potential roles in acquisition, maintenance, and cue-induced relapse; see e.g. Tiffany and Drobes 1990; Gawin 1991; O'Brien *et al.* 1998) as it represents a mechanism by which uncontrolled (noncontingent) stimuli can radically affect goal-directed responding.

Thus, an ostensibly simple behaviour — lever-pressing in rats — is influenced by many dissociable psychological processes. Understanding of these processes is by no means complete. For example, it is not clear how conditioned reinforcement (in which neutral stimuli paired with primary reward gain affective or motivational value such that animals will work for them: see Mackintosh 1974; Williams 1991; Williams and Dunn 1991; Williams 1994a) relates to the valuation processes described above. Conditioned reinforcers might act in several ways (for example, by gaining instrumental incentive value and also affecting behaviour via PIT). Thus, the fact that multiple psychological processes contribute to a behaviour such as lever-pressing must be borne in mind when considering a complex phenomenon such as responding for delayed reinforcement.

3. Mechanisms by which delayed reinforcement can affect instrumental 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 a stimulus–response (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 and Gleeson 1990; Dickinson *et al.* 1992). Interestingly, in several 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, making the ‘underlying’ response–reinforcement delay gradient function unclear. The importance

of this effect was first shown by Grice (1948), who trained rats on a visual discrimination task with delayed reinforcement. The rats were given a choice between a white and a black alley; both led to a grey goal box, but only one (reached from the white alley) contained food. A delay to reinforcement was added by inserting two grey alleys of variable length. Grice found that learning was noticeably impaired by as short a delay as 0.5 s, and severely impaired by 5 s, a deficit that could be ameliorated by having more discriminable goal boxes or by forcing the rats to make discriminable motor responses in the black and white start alleys. Grice argued that learning under conditions of delayed reward was due to immediate secondary (conditioned) reinforcement, based on traces of visual or proprioceptive stimuli. Clearly, if the 'primary' delay gradient is a meaningful concept, it is steep; the distinction becomes one of whether the delay gradient applies to direct reinforcement of responses (Hull) or to stimulus–reward association (Grice).

Killeen and Fetterman (1988), in contrast, suggest 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' arises 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 and Garcia (1970); see also Mackintosh (1974, pp. 155-159).

When considering choice between reinforcers when one is delayed, one reason for the failure of a subject to choose the delayed reinforcer may therefore be that a response–reinforcer, stimulus–reinforcer, or stimulus–response association is weaker for the delayed alternative. Additionally, if the delay is bridged by a stimulus acting as a conditioned

reinforcer, then neural interventions known to increase the power of conditioned reinforcers should improve subjects' ability to choose the delayed reward.

4. Delayed reinforcement in choice

4.1. Utility theory as an approach to normal and pathological choice in animals

To analyse choice involving delayed reinforcement, it is natural to attempt to quantify the value of reinforcers to the subject, and utility theory is a way of analysing choice that has explicitly or implicitly underlain many studies using delayed reinforcement. Formal utility theory is based on six axioms defining attributes of preference that perfectly rational agents should possess (von Neumann and Morgenstern 1947) (reviewed by Russell and Norvig 1995). A full statement of these axioms is beyond the scope of this article, but 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 $A>B>C>A$, and initially possesses *A*, then an observer could offer the agent *C* in exchange for *A* plus 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 an agent obeys the axioms of utility theory, 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 $EU(\text{action}|\text{evidence}) = \sum_n p(action \rightarrow outcome_n | evidence) \cdot U(outcome_n)$. 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.

However, this decision-making approach suffers from two particular deficiencies. Firstly, computing the expected utilities takes finite time. In real-world situations it may often be better to make an imperfect decision quickly than *eventually* to make what *would* have been the perfect decision — a difficult problem (Russell and Norvig 1995). Secondly, the MEU principle implies that in identical situations, the same action will always be taken (it is a 'pure' strategy). Once again, this may not be a wise real-world strategy. Game theory (von Neumann and 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), and how randomness is used by animals in decision-making is poorly understood (see Mérö 1998).

Within the framework of utility theory, there are two ways to produce 'pathological' decision-making, such as that contributing to drug addiction. 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 possibility, 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; Heckerman *et al.* 1992; see also Chase *et al.* 1998). The distinction is difficult. As an

illustration, consider a smoker who desires to abstain but then lights a cigarette. Are we to consider the decision flawed or the actual utility of smoking higher than he first 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. Nevertheless, this distinction seems important; if so, a more reductionist approach is required to the way the brain reaches decisions.

4.2. Choice within the brain: 'top-down' and 'bottom-up' approaches

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). Initial attempts involved the calculation of reinforcement efficacy by establishing the relationship between *response rate* and the frequency and amount of reinforcement; however, such studies 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 and Dickinson 1990). Similarly, the mechanisms governing motor aspects of responding are neurally dissociable from motivational mechanisms (see e.g. Robbins and Everitt 1992).

Another approach has been to relate reinforcement efficacy to *choice* behaviour. This literature grew following 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: $R_1 / (R_1 + R_2) = r_1 / (r_1 + r_2)$ 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): $R_1 = kr_1 / (r_1 + r_e)$, 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 and Herrnstein 1976).

This approach has not been without problems. In many circumstances, subjects have been found to 'overmatch' (exhibit relative 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) 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 underlying choice rule (Mackintosh 1974, pp. 192-195; Williams 1994b).

All these theories have a common theme: it is assumed that some value is computed for each alternative behaviour, and a 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 and 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), others, using different techniques, have found that the effects of reinforcer delay and magnitude are not independent (Ito 1985; White and 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.

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. Rather than being fundamental, rationality (and the MEU principle) may represent an ideal that is approximated by a set of heuristic psychological processes implemented in the brain. 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 discussed above, rats possess *at least* two representations of the value of foodstuffs (Dickinson and Balleine 1994), namely hedonic value and the incentive value governing instrumental responding; Pavlovian incentive value is probably a third (see above). A 'bottom-up' 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 may be 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.

4.3. Temporal discounting and impulsive choice

Given these caveats, studies of impulsive choice have produced some highly consistent results regarding the effects of delayed reinforcement in well-defined choice paradigms. In a typical experimental 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 it is assumed 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), using natural reinforcers and intracranial self-stimulation (Grice 1948; Mazur 1987; Mazur *et al.* 1987; Richards *et al.* 1997b), that time discounting is described well by a *hyperbolic* discount function (Figure 1A) or a very similar power law (Grace 1996); see Kacelnik (1997) for a discussion of 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 1B), and such preference reversal is a reliable experimental finding (for references see Bradshaw and Szabadi 1992).

At present, 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. Neuropsychological research along these lines is a young field. However, 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 may contribute to choice, reinforcement and value assessment (reviewed by Cardinal *et al.* in press). Given the importance of impulsive choice in addiction (Poulos *et al.* 1995; Heyman 1996; Bickel *et al.* 1999; Evenden 1999a; Mitchell 1999) and ADHD (Sagvolden *et al.* 1998; Sagvolden and Sergeant 1998), a number of groups have studied the effects on impulsive choice of manipulating neurochemical and neuroanatomical systems implicated in addiction and ADHD; these studies will be reviewed next.

5. Neurochemical studies of impulsive choice

5.1. 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 a 'motor' sense (Soubrié 1986; Evenden 1999b), and that low levels of serotonin metabolites in cerebrospinal fluid are associated with impulsive aggression and violence in humans (Åsberg *et al.* 1976; Linnoila *et al.* 1983; Brown and 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 and 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 5-HT-depleted 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). Conversely, increasing serotonin function with the 5-HT indirect agonist fenfluramine decreases impulsivity (Poulos *et al.* 1996).

5.2. Dopamine (DA)

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 and 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, are due to an abnormally short and steep delay gradient, and that this is due to a hypofunctional dopamine system (Sagvolden *et al.* 1998). Indeed, they go on to suggest nucleus accumbens (Acb) DA as the specific culprit (Sagvolden *et al.* 1998; Sagvolden and Sergeant 1998). Acb DA has long been implicated in aspects of responding for reinforcement, though its role is not yet fully understood (see Cardinal *et al.* in press).

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 and 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 progenitor control rats in a complex pattern that has been attributed to hypofunction of the mesolimbic dopamine system projecting to the Acb (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). 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).

Impulsive choice has been suggested to reflect an alteration in reinforcement processes, namely that delayed reinforcers have lost their effectiveness, and has been suggested to underlie attention-deficit/hyperactivity disorder (ADHD) (Sagvolden *et al.* 1998; Sagvolden and Sergeant 1998). ADHD is amenable to treatment with psychomotor stimulant drugs (Bradley 1937; Solanto 1998), suggesting that they might promote the choice of delayed rewards. However, in laboratory models of impulsive choice, the effects of acute administration of psychostimulants have varied: some studies have found that they promote choice of delayed reinforcers (Sagvolden *et al.* 1992; Richards *et al.* 1999; Wade *et al.* 2000), while others have found the opposite effect (Charrier and Thiébot 1996; Evenden and Ryan 1996), and it has been shown that the same psychostimulant can have opposite effects in different tasks designed to measure impulsivity (Richards *et al.* 1997a).

Intriguingly, in studies of delayed reinforcement, it has been demonstrated that signalled delays generally maintain higher rates of free-operant responding than unsignalled delays (see Lattal 1987 for a review), and signals present during the delay to reinforcement can have an important role in discrete-trials choice (Mazur 1997). A signal or cue that is associated selectively with a reinforcing outcome may become a conditioned reinforcer, and conditioned reinforcement can affect choice behaviour (Williams and Dunn 1991). Since amphetamine-like drugs potentiate the effects of conditioned reinforcers (Hill 1970; Robbins 1976; Robbins 1978; Robbins *et al.* 1983), Cardinal *et al.* (2000) tested the hypothesis that amphetamine promotes the choice of signalled delayed reinforcement by potentiating conditioned reinforcing properties of the cue. Rats were given regular discrete-trial choices between an immediate small reinforcer and a delayed large reinforcer. For one group of rats, illumination of a stimulus light during the delay provided a signal that was unambiguously associated with the large reinforcer (termed the Cue condition; Figure 2).

This design is commonly used to establish stimuli as conditioned reinforcers in delay-of-reinforcement experiments (for reviews see Williams 1994a; Mazur 1997). In the No Cue condition, rats awaited and collected the reinforcers in darkness, with no signal present during the delay. Given that the effect of amphetamine on performance of this task in the absence of differential cues was to increase preference for the small immediate reward (reduced tolerance of delay, Evenden and Ryan 1996), the addition of a conditioned reinforcer would be expected to reduce or reverse this effect. To prevent rats accumulating more food overall by choosing the immediate reward frequently, the time between successive trials was held constant (see Sonuga-Barke *et al.* 1998); thus, choice of the small immediate reinforcer was always suboptimal.

Cardinal *et al.* (2000) found that the effects of amphetamine depended on the effects of the cue, decreasing preference for the large, delayed reinforcer in the No Cue condition, but increasing it in the Cue condition. This finding is consistent with Evenden and Ryan (1996), who used a task equivalent to the No Cue condition of Cardinal *et al.* (2000) and found that amphetamine reduced preference for the delayed reward. It is also consistent with results obtained using an adjusting-amount procedure (Richards *et al.* 1997b) in which a tone was sounded for the duration of the delay, analogous to the Cue condition of Cardinal *et al.* (2000); in this task, amphetamine and the amphetamine analogue methamphetamine increase preference for the larger, delayed reward (Richards *et al.* 1997a; Richards *et al.* 1999; Wade *et al.* 2000). The results of Cardinal *et al.* (2000) therefore support the idea that 'delay discounting' of the efficacy of future rewards is not a unitary process (Ainslie 1975), but rather that the observed phenomenon of discounting arises from several underlying processes of which conditioned reinforcement is one.

6. Neuroanatomical studies of impulsive choice

In contrast to the literature on the neurochemistry of impulsivity, little is known of the neuroanatomical basis of impulsive choice. However, three lines of evidence have suggested the nucleus accumbens (Acb) and its cortical afferents, including the anterior cingulate and medial prefrontal cortices (ACC, mPFC), as candidate structures that may be involved in regulating choice between alternative reinforcers.

First, these structures have been firmly implicated in reinforcement processes. The Acb, once suggested to mediate the reinforcing efficacy of natural and artificial rewards (see Koob 1992) (and also Wise 1981; 1982; 1985; 1994)**,** is now thought not to be necessary for this, but instead to be a key site for the motivational impact of impending rewards (reviewed by Robbins and Everitt 1996; Salamone *et al.* 1997; Everitt *et al.* 1999; Parkinson *et al.* 2000a; Cardinal *et al.* in press). Many of its afferents have also been shown to be involved in reward-related learning, including the ACC (Bussey *et al.* 1997a; Bussey *et al.* 1997b; Parkinson *et al.* 2000b) and mPFC (e.g. Balleine and Dickinson 1998; Richardson and Gratton 1998; Bechara *et al.* 1999; Tzschentke 2000).

Second, these regions are important recipients of dopaminergic and serotonergic afferents (Fallon and Loughlin 1995; Halliday *et al.* 1995), and pharmacological manipulations of dopamine and serotonin systems have been shown to affect impulsive choice in rats, as described above.

Third, abnormalities of these regions have been detected in humans with ADHD, and in animal models of ADHD. Abnormal functioning of prefrontal cortical regions, including medial prefrontal and anterior cingulate cortex, has been observed in ADHD patients (Ernst *et al.* 1998; Bush *et al.* 1999; Rubia *et al.* 1999). In the spontaneously hypertensive rat (SHR), differences in dopamine receptor density and gene expression have been observed within the core and shell regions of the Acb (Papa *et al.* 1996; Carey *et al.* 1998;

Papa *et al.* 1998; Sadile 2000). Abnormalities of dopamine release have been detected in the Acb (de Villiers *et al.* 1995; Russell *et al.* 1998; Russell 2000) and prefrontal cortex (Russell *et al.* 1995), in addition to possible dysfunction in the dorsal striatum and amygdala (Russell *et al.* 1995; Papa *et al.* 2000).

6.1. Role of the nucleus accumbens (Acb)

Using the task described earlier (Figure 2), with no cues present during the delay to avoid the potential confound of conditioned reinforcement, Cardinal *et al.* (2001) examined the effects of excitotoxic lesions of the nucleus accumbens core (AcbC; Figure 3) on rats' ability to choose a delayed reward. Subjects were trained pre-operatively and assigned to matched groups before being tested post-operatively, to avoid any possible effects of the lesion on learning of the task.

AcbC-lesioned subjects exhibited a persistent and profound deficit in subjects' ability to choose a delayed reward, making impulsive choices (Figure 4). This was not due to an inflexible bias away from the lever producing the delayed reinforcer: AcbC-lesioned rats still chose the large reinforcer more frequently at zero delay than at other delays, and removal of the delays resulted in a rapid and significant increase in the rats' preference for the large reinforcer. Thus, the pattern of choice reflected a reduced preference for the large reinforcer when it was delayed, suggesting that delays reduced the effectiveness or value of reinforcers much more in AcbC-lesioned rats than in controls.

In the initial set of post-operative testing sessions, the AcbC-lesioned rats preferred the small reinforcer even at zero delay, avoiding the large reinforcer. This paradoxical finding probably reflects an induced bias away from the lever providing delayed reinforcement (Evenden and Ryan 1996; Cardinal *et al.* 2000). However, the majority of AcbC-lesioned subjects (6/10) showed a consistent preference for the large reinforcer after prolonged training without delays (Figure 5A). This did not overcome the tendency to avoid the lever

previously associated with delayed reinforcement in all lesioned subjects; given that the pre-operative performance of the same animals was equal to that of controls, it is possible that the post-operative experience of delayed reinforcement may have been aversive for AcbC-lesioned rats (or at least, much less preferable than immediate small reinforcement), inducing them to avoid that lever permanently. However, even when sham and AcbC-lesioned subjects were selected who showed near-exclusive preference for the large reinforcer in the absence of delays (Figure 5E), reintroduction of delays caused a dramatic and selective fall in preference for the large, delayed reinforcer in the AcbC-lesioned group (accompanied by a decline in preference at zero delay, Figure 5F–H, which represents generalization of their avoidance of the delayed reinforcer from trial blocks on which delays were present to the first trial block). In summary, these results show that AcbC-lesioned rats are hypersensitive to the effects of delaying reinforcement even when they are clearly able to discriminate the two reinforcers.

The task used by Cardinal *et al.* (2001) does not determine whether AcbC-lesioned rats exhibit altered sensitivity to reinforcer *magnitude* or *delay*. Either abnormality might produce impulsive choice (see Ho *et al.* 1999). However, in their study, the AcbC group did discriminate between the large and the small reinforcer, consistent with the observation that the expectancy of reward magnitude continues to have normal effects upon rats' reaction time following excitotoxic Acb lesions (Brown and Bowman 1995). A large proportion of the AcbC group showed a preference, sometimes absolute, for the large reward when prolonged training was given with no delays, and 5/10 AcbC-lesioned rats met a very stringent criterion of preference for the large reward under these conditions (Figure 5E). These same rats were exquisitely sensitive to delays, preferring the large reinforcer much less than similarly-selected shams when it was delayed.

It is possible that the AcbC group discriminated between the reinforcer magnitudes, but to a lesser extent than normal rats. In this scenario, AcbC-lesioned rats exhibit impulsive

choice because the perceived value of the large reinforcer is abnormally low, and insufficient to overcome the normal effects of delay discounting. Models such as the multiplicative hyperbolic discounting model of Ho *et al.* (1999) have been derived based on behavioural techniques allowing magnitude and delay discounting parameters to be determined independently. Unfortunately, the behavioural technique used by Cardinal *et al.* (2001) cannot be analysed using this model. For example, sham subjects' preferences approached 100% choice of the large reinforcer at zero delays (Figure 5A), whereas in the model of Ho *et al.* (1999), relative preference between a 1-pellet and a 4-pellet reinforcer cannot exceed 80%. The behavioural result comes as no surprise, for it is the well-known phenomenon of maximization on discrete-trial schedules (see Mackintosh 1974, pp. 190-195), but it implies that behaviour in this task cannot be quantified according to the hyperbolic discounting model.

Thus, the task used by Cardinal *et al.* (2001) does not allow the two explanations of impulsive choice (variations in sensitivity to reinforcer magnitude or delay) to be distinguished conclusively. While this may be possible in delay-of-reinforcement choice tasks using indifference-point methodology (Ho *et al.* 1999), there is an alternative. Relative preference for two reinforcers is often inferred from the distribution of responses on concurrent VI schedules of reinforcement (discussed ealier), which, while complex to interpret when delayed reinforcement is used, are simpler to interpret with immediate reinforcement. Such a schedule could be used to assess whether or not AcbC-lesioned rats exhibited relative indifference ('undermatching' compared to shams) between the reinforcers used by Cardinal *et al.* (2001). This would provide evidence for reduced reinforcer magnitude discrimination following AcbC lesions, or for an abnormality of the matching process itself, while normal performance (or overmatching) would make this explanation less likely and therefore support the view that AcbC lesions produce a steeper

delay-of-reinforcement gradient. As yet, published data do not allow this question to be answered.

Finally, an explanation in terms of temporal perception might also be offered for the effects of AcbC lesions. The basal ganglia have been suggested to be a component of an 'internal clock', based on the effects of dopaminergic manipulations on timing tasks (see Gibbon *et al.* 1997). Similarly, forebrain serotonin depletion that affects Acb, among many other structures, impairs timing ability (Morrissey *et al.* 1993; Wogar *et al.* 1993a; Morrissey *et al.* 1994; Al-Zahrani *et al.* 1997), though these impairments sometimes reflect enhanced behavioural switching rather than a true timing deficit (Ho *et al.* 1995; Al-Zahrani *et al.* 1996; Al-Ruwaitea *et al.* 1997a); see Al-Ruwaitea *et al.* (1997b) for a review. Applying the distinctions of Killeen and Fetterman (1988) to the task of Cardinal *et al.* (2001) (Figure 2), a lesion that increased the speed of an 'internal clock' might affect choice *prospectively* (i.e. the lesioned subject perceives itself to be at a later time-point in the session than it actually is, hastening the within-session shift towards the lever producing immediate reinforcement), or might affect *retrospective* choice (i.e. the lesioned subject experiences a given delay as longer than it remembered, causing a decrease in its preference for the lever producing delayed reinforcement). Again, there is at present no evidence to address the question of whether excitotoxic AcbC lesions affect behavioural timing.

At the least, the findings of Cardinal *et al.* (2001) show that the Acb contributes significantly to animals' ability to choose a delayed reward. If further experiments show that it does so specifically by maintaining the value of a reinforcer over a delay, a new avenue of inquiry into Acb function might open up. It has previously been shown in primates that neuronal activity related to the expectation of reward across a delay can be found in the ventral striatum (Schultz *et al.* 1992; Schultz *et al.* 1995; Schultz *et al.* 1998; Schultz *et al.* 2000); such activity is a candidate representation of the goals of activity

(Schultz *et al.* 2000). Additionally, striatal neurons may respond to past events, maintaining a form of memory that might assist the association of past acts with reinforcement (Schultz *et al.* 1995). These findings represent important data on the forms of information that the AcbC may use to promote actions leading to delayed rewards, and a future challenge will be to discover the manner in which these neural signals influence overt behaviour, and the psychological processes they govern. Given the involvement of the Acb in aversive motivation (see Salamone 1994; Parkinson *et al.* 1999), it will also be important to determine whether lesions of Acb induce impulsive choice in an aversive context, impairing the ability to choose a small immediate penalty in preference to a large delayed penalty.

6.2. Role of the anterior cingulate cortex (ACC)

Cardinal *et al.* (2001) found that excitotoxic lesions of the ACC (Figure 3) had no effect on choice in the same task, establishing that the ACC is not required for rats to choose a delayed reinforcer (Figure 4). Moreover, ACC-lesioned rats remained equally sensitive to unexpected removal of the delays in this task (Figure 4), suggesting that their choices were no more inflexible or 'habitual' than those of shams. This finding stands in apparent contrast to previous reports of motor impulsivity or disinhibited responding in ACC-lesioned rats. For example, such rats have been found to over-respond to unrewarded stimuli (Bussey *et al.* 1997a; Parkinson *et al.* 2000b), and to respond prematurely in situations where they are required to wait (Muir *et al.* 1996). However, such a dissociation is not in itself unexpected, as motor impulsivity and impulsive choice have been dissociated before ('execution' and 'outcome' impulsivity) (Evenden 1999b). Thus, these results suggest that despite findings of ACC abnormalities in disorders of impulsivity (e.g. Bush *et al.* 1999), ACC dysfunction is not an important contributor to impulsive choice.

6.3. Role of the medial prefrontal cortex (mPFC)

In the task used by Cardinal *et al.* (2001), lesions of the mPFC 'flattened' the within-session shift from the large to the small reward; the mean preference for the large reward was *below* that of shams at zero delay, but *above* that of shams at the maximum delay (Figure 4F). There is no obvious explanation for this effect within theories of choice of delayed reinforcement; it seems clear that the mPFC lesion resulted in some form of insensitivity to the contingencies or stimuli present in the task.

Given that Balleine and Dickinson (1998) demonstrated that lesions encompassing prelimbic cortex impaired rats' sensitivity to instrumental contingencies, it might be that a failure of contingency perception was responsible for performance of mPFC-lesioned rats in the present task. However, these rats were just as sensitive as controls to the unexpected removal of all delays; their responding was not inflexible, as might have been expected according to this account.

A more plausible interpretation is that mPFC lesions disrupted the control over behaviour by the passage of time in each session. There is strong evidence that normal rats learn a session-wide temporal discrimination in this task, and that this temporal discriminative stimulus comes to control responding — in particular the tendency to shift from the large to the small reward as the session progresses (Cardinal *et al.* 2000). Disruption of such temporal stimulus control might be expected to produce a flattening of the within-session shift of the kind seen. Indeed, aspirative lesions of the mPFC have previously been shown to induce a general deficit in timing ability in rats (Dietrich and Allen 1998); lesioned subjects showed a temporal discrimination function that was less steep than normal in the peak procedure, an operant task that assesses the ability to time a discriminative stimulus (Catania 1970; Roberts 1981).

6.4. Role of the orbitofrontal cortex (OFC)

Finally, Mobini *et al.* (2002) have recently studied the effects of lesions encompassing the OFC on a task very similar to those of Evenden and Ryan (1996) and Cardinal *et al.* (2001). The OFC is a prefrontal cortical region strongly implicated in the assessment of reward value that projects to the AcbC (see Cardinal *et al.* in press). Intriguingly, OFC lesions induced impulsive choice in a manner very similar to AcbC lesions (and, as for AcbC lesions, it is not known whether the impulsive choice was as a result of altered sensitivity to reinforcer magnitude or delay; see above). Although the lesions of Mobini *et al.* (2002) damaged the prelimbic cortex in addition to the OFC, their hypothesis that OFC damage was responsible for the behavioural effect is strengthened by the finding of Cardinal *et al.* (2001) that mPFC lesions encompassing the prelimbic cortex did not induce impulsive choice.

7. Conclusions

7.1. Theories of learning and choice with delayed reward

Two broad approaches to choice behaviour will be summarized, and a synthesis offered.

Model 1: informed choice. According to this model, subjects make prospective choices between alternatives based on full knowledge of the response–outcome contingencies and of the value of each outcome. These choices represent goal-directed actions. Subjects' sensitivity to delay in choice tasks is therefore a consequence of temporal discounting of the perceived (prospective) value of the delayed reward.

This model is necessarily applicable only to fully-trained subjects — subjects who have learned the instrumental contingencies. It may be particularly applicable when humans are offered explicit hypothetical choices: 'would you prefer $800 now, or $1000 in a year?' (Rachlin *et al.* 1991; Myerson and Green 1995). As the contingencies cannot be offered 'pre-packaged' to experimental animals through language, such subjects must be trained

through direct experience of the rewards in the experimental situation. This introduces the complication that delays to reinforcement can affect operant and discrimination learning, so care is typically taken by experimenters to ensure subjects are 'well trained'. Slow acquisition of delay sensitivity must be attributed to difficulties in learning the instrumental contingencies across a delay and / or learning the appropriate incentive value of delayed reward through experience of waiting. In tasks where the delay is systematically and predictably varied (e.g. Evenden and Ryan 1996; Cardinal *et al.* 2000), learning may also be slowed by the requirement to learn discriminative stimuli predicting the delay contingency currently in force. Thus, this model is inherently an incomplete description of the effects of delayed reinforcement, as it does not deal with the effects of delays on learning.

Model 2: differential associative response strength. According to an extreme form of this model, based on simple S–R theory (Thorndike 1911; Grindley 1932; Guthrie 1935; Hull 1943), rats' choice behaviour reflects differential reinforcement of stimulus–response habits. The change in associative strength is some function of reward magnitude multiplied by the time-discounted 'trace strength' of the preceding response. Choice is determined by a process of competition between the available responses (e.g. the principles of matching: Herrnstein 1970; de Villiers and Herrnstein 1976). Choice is therefore 'retrospective' in a sense, as preference for a particular alternative depends upon prior experience of that alternative, and time discounting reflects the decay of the traces available to be associated with reward. A similar model, after Grice (1948), may be constructed in which animals respond for immediate conditioned reinforcement (by goal-directed behaviour or S–R habit) and the acquisition of associations between a chain of stimuli and eventual reward accounts for the observed phenomenon of temporal discounting, by similar mechanisms.

The S–R view accounts for some of the theoretical appeal of exponential temporal discounting models. In exponential decay, at any one moment in time the trace strength of

a response follows directly from the trace strength at the previous instant (if x_t is the trace strength at time t and A is the starting value, then $x_t = \mathrm{A}e^{-kt}$ and $x_{t+1} = e^{-k}x_t$). In contrast, in the hyperbolic discounting model and all others in which preference reversal occurs, the strength of the trace at any one moment cannot be calculated in such a manner: information about the absolute time since the response must be available. (This is clearly illustrated by the preference reversal graphs in Figure 1B; if two such decay curves cross, then an observer travelling along one curve cannot know at the crossover point whether its own curve is the recent, rapidly-decaying trace, or the older, slowly-decaying trace, without further information — namely the time since the response, or its starting strength.) This process does not model 'mnemonic decay' in any clear way. Thus, the empirical observation of *hyperbolic* discounting specifies the information that must be available to the subject at any one moment in time; in the context of 'retrospective' choice, this constrains the possible underlying psychological mechanisms, and there is no obvious candidate within the S–R model.

Furthermore, while S–R models can account for effects of delays on learning as well as choice, they do not take into account the fact that goal-directed actions contribute to choice in rats (Dickinson 1994) and would clearly not provide a satisfactory account of human choice (cf. Ainslie 1975; Rachlin *et al.* 1991; Myerson and Green 1995).

Model 3: composite. A multifactorial model is therefore suggested, based on that of Dickinson (1994). The 'response strength' of any behaviour is governed by (1) goal-directed action (Dickinson and Balleine 1994), in which knowledge of the instrumental contingency combines with the incentive value of the expected outcome; (2) stimulus–response habits, which gain strength slowly with the number of reinforcers presented (Dickinson *et al.* 1995); and (3) Pavlovian–instrumental transfer, mediated by the Pavlovian association between contextual, discriminative, or other conditioned stimuli and the outcome of the instrumental action. Ordinarily, behaviour conforming to the matching

law and to hyperbolic temporal discounting is seen as a product of these processes. Delayed reinforcement may act (a) to impair learning of the instrumental contingency (Dickinson *et al.* 1992); (b) to reduce the incentive value of the delayed reward, as speculated by many models; (c) to reduce the reinforcement of stimulus–response habits; and (d) to reduce the Pavlovian association between stimuli present at the moment of action and the ultimate reinforcer.

This model, though qualitative, makes several predictions. Firstly, manipulations of components of this composite behaviour should affect choice. For example, manipulations of the association between cues immediately consequent on choice and the outcome (e.g. the presence or absence of a cue bridging the delay) should affect choice independently of the actual delay to reinforcement, a prediction not made by Kacelnik's (1997) normative model of hyperbolic discounting, but one that has experimental support (Mazur 1997; Cardinal *et al.* 2000). Secondly, pharmacological and neural manipulations known to dissociate these processes should also be capable of affecting choice (Cardinal *et al.* 2000).

This model is obviously compatible with mathematical models of temporal discounting, but interprets the discount function as the sum of the contributions of several processes operating in any one situation. Similar composite models have been offered before (a casual example is Pinker 1997, pp. 395-396), though with a different decomposition of the processes contributing to choice (compare, for example, distinct contributions of conditioned and primary reinforcement to response strength: Killeen and Fetterman 1988, pp. 287-289). One interesting challenge may be to establish what processes contribute most significantly to choice of a reinforcer at different delays. Consider an obvious hypothesis: instrumental incentive value in the rat depends upon declarative knowledge, as discussed above, and in this way is analogous to human hypothetical choices. Thus it may be that when reward is extremely delayed (as in some human experiments), only instrumental incentive value is important (as delay $d \rightarrow \infty$, total value $V \rightarrow V_{\text{instrumental}}$). When a dieting

human calmly decides to abstain from chocolate cake and the dessert trolley is then pushed under his nose, it would not be expected from the rat literature (see Cardinal *et al.* in press) that the instrumental incentive value of chocolate cake suddenly changes — after all, the subject's underlying motivational state of hunger (or lack of it) has not altered. However, alternative, possibly Pavlovian motivational processes may create an extra boost to the value of the cake (observed as a tendency to choose the cake), as the cake is now immediately available (as $d \rightarrow 0$, $V_{\text{cake–other}}$ increases dramatically). The net value function ($V_{\text{cake}} = V_{\text{cake–instrumental}} + V_{\text{cake–other}}$, versus $V_{\text{weight loss}}$) could then exhibit preference reversal, leading our diner to succumb and choose the immediate reinforcer. This illustrates but one possible scenario. Nevertheless, if different processes do contribute at different delays, there would be important implications for our understanding of individual differences in impulsive choice.

7.2. Neural basis of choosing delayed reinforcers

As described above, there is now direct evidence that the Acb is involved in the pathogenesis of impulsive choice: the integrity of the Acb is critical for animals to tolerate delays to appetitive reinforcement (Cardinal *et al.* 2001). In addition to providing neuroanatomical insight into the normal process through which delayed reinforcement affects behaviour, this finding suggests a mechanism by which Acb dysfunction may contribute to addiction, ADHD, and other impulse control disorders.

As lesions of the ACC and mPFC, two afferents to the AcbC, did not impair rats' capacity to choose a delayed reward (Cardinal *et al.* 2001) it remains to be established which afferents to the AcbC contribute to its ability to promote the choice of delayed rewards, and through what efferent pathways it does this. Obvious afferent structures that may provide specific information concerning reinforcer value to the Acb are the basolateral amygdala and the OFC, both implicated in the assessment of reward value (see Cardinal *et*

al. in press). The OFC may also be an important efferent target of information travelling through the Acb, as this 'limbic loop' of the basal ganglia projects back (through the ventral pallidum) to medial OFC (Alexander *et al.* 1986). In support of the conjecture that an OFC–AcbC circuit may play an important role in animals' ability to choose a delayed reinforcer (Cardinal *et al.* 2001), Mobini *et al.* (2002) recently found that OFC lesions induce impulsive choice in rats in a manner very similar to AcbC lesions. If this hypothesis is correct, then a 'disconnection' lesion (a lesion of the OFC in one hemisphere and the AcbC in the other) should impair subjects' ability to choose delayed reinforcement. In addition, it remains to be seen whether the nucleus accumbens shell (AcbSh) also plays a role in the choice of delayed rewards. This is another interesting target of investigation, given the abnormalities of dopamine receptor function detected in the AcbSh of the SHR (Papa *et al.* 1996; Carey *et al.* 1998; Papa *et al.* 1998; Sadile 2000).

Finally, the limbic corticostriatal circuit of which the Acb is a part may not be the only system involved in delayed reinforcement. In principle, any structure that represents future reinforcers across a delay may contribute to their choice, and exert conditioned reinforcing effects on current behaviour, while any structure that maintains a 'memory trace' of responses across a delay may support the reinforcement of those responses. The ventral striatum and OFC exhibit such activity (Schultz *et al.* 1995; Schultz *et al.* 1998; Schultz *et al.* 2000), but so do other structures including the dorsal striatum (e.g. Schultz *et al.* 1995), implicated in the reinforcement of stimulus–response habits (Mishkin *et al.* 1984; Robbins and Everitt 1992; Packard and McGaugh 1996; White 1997; see Parkinson *et al.* 2000a).

7.3. Application to addiction and other disorders of impulsivity

The observation that AcbC damage can induce impulsive choice has implications for the understanding of ADHD and drug addiction, two clinical disorders in which impulsive choice is a factor.

AcbC-lesioned animals exhibited two signs of ADHD (APA 1994; Sagvolden and Sergeant 1998): locomotor hyperactivity and impulsive choice. However, such animals do not exhibit attentional deficits: neither 6-OHDA-induced dopamine depletion of the Acb (Cole and Robbins 1989) nor excitotoxic lesions of the AcbC (A. Christakou, T.W. Robbins, and B.J. Everitt, unpublished data) affect accuracy in tests of visuospatial attentional function. Thus, AcbC-lesioned rats may represent an animal model of the hyperactive–impulsive subtype of ADHD (APA 1994). Interventional neuroanatomical studies of impulsive choice are clearly important for the understanding of the pathogenesis of ADHD, for they allow a causal role to be established between dysfunction of a brain region and impulsive choice. Damage to two other structures observed to be abnormal in ADHD — the ACC and mPFC — did not induce impulsive choice; abnormalities of structure or function observed in these regions in ADHD brains may therefore be responsible for other features of the disorder (such as motoric disinhibition), or these regions may have altered as a consequence of a disease process beginning elsewhere. A clearer understanding of the neurochemical and neuroanatomical basis of the symptoms and signs of ADHD may lead to more effective therapy.

The same considerations apply to drug addiction, in which impulsive choice plays a prominent role in maintaining the selection of drugs of abuse in favour of other, longer-term rewards. Furthermore, if the suggestion that Pavlovian (cue-induced) motivational processes contribute to preference reversal effects and to addiction is correct (see above and Gjelsvik, this volume), then the role of the AcbC is doubly important (see Cardinal *et al.* in press), for PIT requires the AcbC (Hall *et al.* 2001); noncontingent CSs elevate AcbC DA levels (Bassareo and DiChiara 1999; Ito *et al.* 2000), DA antagonists block PIT (Dickinson *et al.* 2000), and enhancement of Acb DA function boosts PIT (Wyvell and Berridge 2000). The process of addiction is complicated further by the ability of drugs of abuse (including opiates, ethanol, and psychostimulants) to produce chronic

neuroadaptations in brain regions including the Acb (see Koob *et al.* 1998). Addictive drugs may be unique among reinforcers in producing sensitization, the phenomenon by which repeated drug administration leads to an enhanced response to the drug (Robinson and Berridge 1993; Altman *et al.* 1996; Kalivas *et al.* 1998); psychostimulant sensitization enhances the sensitivity of the Acb to DA stimulation (Cador *et al.* 1995), and enhances PIT subsequently (Wyvell and Berridge 2001); chronic methamphetamine has also been shown to increase impulsive choice (Richards *et al.* 1999). One mechanism contributing to addiction may therefore be the ability of drugs of abuse to induce damage or dysfunction in the AcbC, further promoting subsequent impulsive choice and future drug-taking.

Acknowledgements

This work was supported by a Wellcome Trust Programme Grant and conducted within the MRC Co-operative for Brain, Behaviour and Neuropsychiatry. RNC was in receipt of a research studentship from the UK Medical Research Council and a James Baird award from the University of Cambridge School of Clinical Medicine.

Figure legends

Figure 1. (A) Hyperbolic discounting, governed by the equation $Value = magnitude / (1 + K \cdot delay)$. Large values of K give the steepest curve. **(B)** Preference reversal. 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).

Figure 2. Delayed reinforcement choice task (Cardinal *et al.* 2000; Cardinal *et al.* 2001), based on Evenden and Ryan (1996). The figure shows the format of a single trial; trials began at 100-s intervals. Sessions lasted 100 min and consisted of 5 blocks, each comprising two trials on which only one lever was presented (one trial for each lever, in randomized order) followed by ten choice trials. The delay to the large reinforcer was varied systematically across the session: delays for each block were 0, 10, 20, 40, and 60 s respectively. In the Cue condition, a stimulus light was illuminated during the delay to the large reinforcer; this was absent in the No Cue condition.

Figure 3. Schematic of lesions of the mPFC (left), ACC (middle), and AcbC (right). Black shading indicates the extent of neuronal loss common to all subjects; grey indicates the

area lesioned in at least one subject. Coronal sections are +4.7 through +1.7 mm (mPFC), +2.7 mm through –1.3 mm (ACC), and +2.7 through +0.48 mm (AcbC) relative to bregma. Outlines are taken from Paxinos and Watson (1998).

Figure 4. Effect of lesions of the AcbC (top), ACC (middle), or mPFC (bottom) on choice of delayed reward (● lesioned group; Δ corresponding sham group; error bars, SEM). The 'no cue' condition (see Figure 2) was used throughout. **A–C** shows the pattern of choice in the last 3 sessions preceding surgery; corresponding sham/lesion groups were matched for performance. Subjects' preference for the large reinforcer declined with delay, as is typical for trained subjects performing this task (Evenden and Ryan 1996; Cardinal *et al.* 2000). **D–F** illustrates choice in the first 7 post-operative sessions. The AcbC-lesioned group was markedly impaired (** $p < .01$), choosing the delayed reinforcer significantly less often than shams at every delay, including zero. However, both groups still exhibited a within-session shift in preference. ACC lesions had no effect on choice. The mPFC-lesioned subjects exhibited a 'flatter' within-session preference shift than shams (# $p < .05$, group × delay interaction). **G–I** illustrates the effects of omitting all delays in alternating sessions (●/○, lesioned/sham groups with delays; ▲/Δ, lesioned/sham groups without delays; error bars, SED for the three-way interaction). All groups remained sensitive to the contingencies. Delay removal increased both the sham- and AcbC-lesioned groups' preference for the larger reward; ACC- and mPFC-lesioned rats were also as sensitive to removal of the delays as shams. Figure reproduced with permission from Cardinal *et al.* (2001).

Figure 5. A illustrates the preference of AcbC-lesioned rats following extended training in the absence of any delays (a further six sessions after completion of other behavioural tests; ● AcbC-lesioned group; Δ shams; error bars, SEM). **B–D** show performance over consecutive blocks of sessions upon the reintroduction of delays (* $p < .05$, ** $p < .01$,

difference from shams). Panels **E–H** show data from the same sessions as **A–D,** but only include data from those rats selected for ≥90% preference for the large reinforcer in every trial block on the last day of training with no delays. The sham and lesioned groups were therefore matched in **E. F–H** show that despite this matching, preference for the large reinforcer in the AcbC group collapsed upon reintroduction of the delays. As these data exhibit significant heterogeneity of variance, the highly conservative correction of Box (1954) was applied (see Howell 1997, pp. 322/457/464); * $p < .05$ for the corrected between-group difference.

References

Adams CD (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **34,** 77-98.

Ainslie G (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, **82,** 463-496.

Alexander GE, DeLong MR and Strick PL (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, **9,** 357-381.

Al-Ruwaitea AS, Al-Zahrani SS, Ho MY, Bradshaw CM and Szabadi E (1997a). Effect of central 5-hydroxytryptamine depletion on performance in the "time-left" procedure: further evidence for a role of the 5-hydroxytryptaminergic pathways in behavioural "switching". *Psychopharmacology*, **134,** 179-186.

Al-Ruwaitea ASA, Al-Zahrani SSA, Ho MY, Bradshaw CM and Szabadi E (1997b). 5-Hydroxytryptamine and interval timing. In CM Bradshaw and E Szabadi, eds. *Time and Behaviour: Psychological and Neurobehavioural Analyses*, pp. 517-570. Elsevier, Amsterdam.

Altman J, Everitt BJ, Glautier S*, et al.* (1996). The biological, social and clinical bases of drug addiction: commentary and debate. *Psychopharmacology*, **125,** 285-345.

Al-Zahrani SS, Ho MY, Velazquez Martinez DN, Lopez Cabrera M, Bradshaw CM and Szabadi E (1996). Effect of destruction of the 5-hydroxytryptaminergic pathways on behavioural timing and "switching" in a free-operant psychophysical procedure. *Psychopharmacology*, **127,** 346-352.

Al-Zahrani SSA, Ho MY, Al-Ruwaitea ASA, Bradshaw CM and Szabadi E (1997). Effect of destruction of the 5-hydroxytryptaminergic pathways on temporal memory: Quantitative analysis with a delayed interval bisection task. *Psychopharmacology*, **129,** 48-55.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders, version IV (DSM-IV)*, American Psychiatric Association, Washington DC.

Aristotle (350 BC / 1925). *Nicomachean Ethics [translated by W.D. Ross]*, Clarendon Press, Oxford.

Åsberg M, Träskman L and Thorén P (1976). 5-HIAA in the cerebrospinal fluid: a biochemical suicide predictor. *Archives of General Psychiatry*, **33,** 1193-1197.

Balleine BW and Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, **37,** 407-419.

Bassareo V and DiChiara G (1999). Differential responsiveness of dopamine transmission to food-stimuli in nucleus accumbens shell/core compartments. *Neuroscience*, **89,** 637-641.

Baum WM (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, **22,** 231-242.

Baum WM (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, **32,** 269-281.

Baum WM and Rachlin HC (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, **12,** 861-874.

Bechara A, Damasio H, Damasio AR and Lee GP (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, **19,** 5473-5481.

Belke TW (1998). Qualitatively different reinforcers and parameters of Herrnstein's (1970) response-strength equation. *Animal Learning & Behavior*, **26,** 235-242.

Bickel WK, Odum AL and Madden GJ (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, **146,** 447-454.

Bizot J, Le Bihan C, Puech AJ, Hamon M and Thiébot M (1999). Serotonin and tolerance to delay of reward in rats. *Psychopharmacology*, **146,** 400-412.

Box GEP (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics*, **25,** 290-302.

Bradley C (1937). The behavior of children receiving Benzedrine. *American Journal of Psychiatry*, **94,** 577-585.

Bradshaw CM and Szabadi E (1992). Choice between delayed reinforcers in a discrete-trials schedule - the effect of deprivation level. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **44B,** 1-16.

Brown GL and Linnoila M (1990). CSF serotonin metabolite (5HIAA) studies in depression, impulsivity and violence. *Journal of Clinical Psychiatry*, **51 (supplement 4),** 31-41.

Brown VJ and Bowman EM (1995). Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. *European Journal of Neuroscience*, **7,** 2479-2485.

Bush G, Frazier JA, Rauch SL*, et al.* (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the Counting Stroop. *Biological Psychiatry*, **45,** 1542-1552.

Bussey TJ, Everitt BJ and Robbins TW (1997a). Dissociable effects of cingulate and medial frontal cortex lesions on stimulus-reward learning using a novel Pavlovian autoshaping procedure for the rat: implications for the neurobiology of emotion. *Behavioral Neuroscience*, **111,** 908-919.

Bussey TJ, Muir JL, Everitt BJ and Robbins TW (1997b). Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behavioral Neuroscience*, **111,** 920-936.

Cador M, Bjijou Y and Stinus L (1995). Evidence of a complete independence of the neurobiological substrates for the induction and expression of behavioral sensitization to amphetamine. *Neuroscience*, **65,** 385-395.

Cardinal RN, Parkinson JA, Hall J and Everitt BJ (in press). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews.*

Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW and Everitt BJ (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, **292,** 2499-2501.

Cardinal RN, Robbins TW and Everitt BJ (2000). The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology*, **152,** 362-375.

Carey MP, Diewald LM, Esposito FJ*, et al.* (1998). Differential distribution, affinity and plasticity of dopamine D-1 and D-2 receptors in the target sites of the mesolimbic system in an animal model of ADHD. *Behavioural Brain Research*, **94,** 173-185.

Catania AC (1970). Reinforcement schedules and psychophysical judgment: A study of some temporal properties of behavior. In WN Schoenfeld, ed. *The theory of reinforcement schedules*, pp. 1-42. Appleton-Century-Crofts, New York.

Charrier D and Thiébot MH (1996). Effects of psychotropic drugs on rat responding in an operant paradigm involving choice between delayed reinforcers. *Pharmacology, Biochemistry and Behavior*, **54,** 149-157.

Chase VM, Hertwig R and Gigerenzer G (1998). Visions of rationality. *Trends in Cognitive Sciences*, **2,** 206-214.

Cole BJ and Robbins TW (1989). Effects of 6-hydroxydopamine lesions of the nucleus accumbens septi on performance of a 5-choice serial reaction time task in rats: implications for theories of selective attention and arousal. *Behavioural Brain Research*, **33,** 165-179.

Dawson GR and Dickinson A (1990). Performance on ratio and interval schedules with matched reinforcement rates. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **42,** 225-239.

de Villiers AS, Russell VA, Sagvolden T, Searson A, Jaffer A and Taljaard JJ (1995). Alpha 2-adrenoceptor mediated inhibition of [3H]dopamine release from nucleus accumbens slices and monoamine levels in a rat model for attention-deficit hyperactivity disorder. *Neurochemical Research*, **20,** 427-433.

de Villiers PA and Herrnstein RJ (1976). Toward a law of response strength. *Psychological Bulletin*, **83,** 1131-1153.

Dickinson A (1985). Actions and habits - the development of behavioral autonomy. *Philosophical Transactions of the Royal Society of London, Series B - Biological Sciences*, **308,** 67-78.

Dickinson A (1986). Re-examination of the role of the instrumental contingency in the sodium-appetite irrelevant incentive effect. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **38,** 161-172.

Dickinson A (1994). Instrumental conditioning. In NJ Mackintosh, ed. *Animal Learning and Cognition*, pp. 45-79. Academic Press, San Diego.

Dickinson A and Balleine B (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, **22,** 1-18.

Dickinson A, Balleine B, Watt A, Gonzalez F and Boakes RA (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior*, **23,** 197-206.

Dickinson A and Dawson GR (1987a). Pavlovian processes in the motivational control of instrumental performance. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **39,** 201-213.

Dickinson A and Dawson GR (1987b). The role of the instrumental contingency in the motivational control of performance. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **39,** 77-93.

Dickinson A, Nicholas DJ and Adams CD (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **35,** 35-51.

Dickinson A, Smith J and Mirenowicz J (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral Neuroscience*, **114,** 468-483.

Dickinson A, Watt A and Griffiths WJH (1992). Free-operant acquisition with delayed reinforcement. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology*, **45,** 241-258.

Dietrich A and Allen JD (1998). Functional dissociation of the prefrontal cortex and the hippocampus in timing behavior. *Behavioral Neuroscience*, **112,** 1043-1047.

Ernst M, Zametkin AJ, Matochik JA, Jons PH and Cohen RM (1998). DOPA decarboxylase activity in attention deficit hyperactivity disorder adults. A [fluorine-18]fluorodopa positron emission tomographic study. *Journal of Neuroscience*, **18,** 5901-5907.

Estes WK (1948). Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *Journal of Experimental Psychology*, **38,** 173-177.

Evenden JL (1998). Serotonergic and steroidal influences on impulsive behaviour in rats. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine*, **764**.

Evenden JL (1999a). Impulsivity: a discussion of clinical and experimental findings. *Journal of Psychopharmacology*, **13,** 180-192.

Evenden JL (1999b). Varieties of impulsivity. *Psychopharmacology*, **146,** 348-361.

Evenden JL and Meyerson B (1999). The behavior of spontaneously hypertensive and Wistar Kyoto rats under a paced fixed consecutive number schedule of reinforcement. *Pharmacology, Biochemistry and Behavior*, **63,** 71-82.

Evenden JL and Ryan CN (1996). The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology*, **128,** 161-170.

Everitt BJ, Parkinson JA, Olmstead MC, Arroyo M, Robledo P and Robbins TW (1999). Associative processes in addiction and reward: the role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences*, **877,** 412-438.

Fallon JH and Loughlin SE (1995). Substantia nigra. In G Paxinos, ed. *The Rat Nervous System*, pp. 215-237. Academic Press, London.

Feldman RS, Meyer JS and Quenzer LF (1997). *Principles of neuropsychopharmacology*, Sinauer, Sunderland, Massachusetts.

Gawin FH (1991). Cocaine addiction: psychology and neurophysiology. *Science*, **251,** 1580-1586.

Gibbon J, Malapani C, Dale CL and Gallistel C (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology*, **7,** 170-184.

Grace RC (1996). Choice between fixed and variable delays to reinforcement in the adjusting-delay procedure and concurrent chains. *Journal of Experimental Psychology: Animal Behavior Processes*, **22,** 362-383.

Grice GR (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology*, **38,** 1-16.

Grindley GC (1932). The formation of a simple habit in guinea pigs. *British Journal of Psychology*, **23,** 127-147.

Guthrie ER (1935). *The psychology of learning*, Harper, New York.

Hall J, Parkinson JA, Connor TM, Dickinson A and Everitt BJ (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *European Journal of Neuroscience*, **13,** 1984-1992.

Halliday G, Harding A and Paxinos G (1995). Serotonin and tachykinin systems. In G Paxinos, ed. *The Rat Nervous System*, pp. 929-974. Academic Press, London.

Heckerman DE, Horvitz EJ and Nathwani BN (1992). Toward normative expert systems: Part I. The Pathfinder project. *Methods of Information in Medicine*, **31,** 90-105.

Herrnstein RJ (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4,** 267-272.

Herrnstein RJ (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, **13,** 243-266.

Heyman GM (1996). Resolving the contradictions of addiction. *Behavioral and Brain Sciences*, **19,** 561-610.

Hill RT (1970). Facilitation of conditioned reinforcement as a mechanism of psychomotor stimulation. In E Costa and S Garattini, eds. *International Symposium on Amphetamines and Related Compounds*, pp. 781-795. Raven Press, New York.

Ho M, Mobini S, Chiang T, Bradshaw CM and Szabadi E (1999). Theory and method in the quantitative analysis of "impulsive choice" behaviour: implications for psychopharmacology. *Psychopharmacology*, **146,** 362-372.

Ho MY, Al-Zahrani SS, Velazquez Martinez DN, Lopez Cabrera M, Bradshaw CM and Szabadi E (1995). The role of the ascending 5-hydroxytryptaminergic pathways in timing behaviour: further observations with the interval bisection task. *Psychopharmacology*, **120,** 213-219.

Howell DC (1997). *Statistical Methods for Psychology*. Fourth edition, Wadsworth, Belmont, California.

Hull CL (1932). The goal gradient hypothesis and maze learning. *Psychological Review*, **39,** 25-43.

Hull CL (1943). *Principles of behavior*, Appleton-Century-Crofts, New York.

Ito M (1985). Choice and amount of reinforcement in rats. *Learning & Motivation*, **16,** 95-108.

Ito R, Dalley JW, Howes SR, Robbins TW and Everitt BJ (2000). Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. *Journal of Neuroscience*, **20,** 7489-7495.

Kacelnik A (1997). Normative and descriptive models of decision making: time discounting and risk sensitivity*Characterizing human psychological adaptations (Ciba Foundation Symposium 208)*, pp. 51-70. Wiley, Chichester.

Kahneman D, Slovic P and Tversky A, Eds. (1982). *Judgement Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Kalivas PW, Pierce RC, Cornish J and Sorg BA (1998). A role for sensitization in craving and relapse in cocaine addiction. *Journal of Psychopharmacology*, **12,** 49-53.

Killeen PR (1972). The matching law. *Journal of the Experimental Analysis of Behavior*, **17,** 489-495.

Killeen PR and Fetterman JG (1988). A behavioral theory of timing. *Psychological Review*, **95,** 274-295.

Koob GF (1992). Dopamine, addiction and reward. *Seminars in the Neurosciences*, **4,** 139-148.

Koob GF, Sanna PP and Bloom FE (1998). Neuroscience of addiction. *Neuron*, **21,** 467-476.

Lattal KA (1987). Considerations in the experimental analysis of reinforcement delay. In ML Commons, JE Mazur, JA Nevin and H Rachlin, eds. *Quantitative Analyses of Behavior: V. The Effect of Delay and of Intervening Events on Reinforcement Value*, pp. 107-123. Lawrence Erlbaum, Hillsdale, New Jersey.

Lattal KA and Gleeson S (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, **16,** 27-39.

Linnoila M, Virkkunen M, George T and Higley D (1993). Impulse control disorders. *International Clinical Psychopharmacology*, **8 (Supplement 1),** 53-56.

Linnoila M, Virkkunen M, Scheinin M, Nuutila A, Rimon R and Goodwin FK (1983). Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sciences*, **33,** 2609-2614.

Lovibond PF (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, **9,** 225-247.

Mackintosh NJ (1974). *The Psychology of Animal Learning*, Academic Press, London.

Mazur JE (1987). An adjusting procedure for studying delayed reinforcement. In ML Commons, JE Mazur, JA Nevin and H Rachlin, eds. *Quantitative Analyses of Behavior: V. The Effect of Delay and of Intervening Events on Reinforcement Value*, pp. 55-73. Lawrence Erlbaum, Hillsdale, New Jersey.

Mazur JE (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning & Behavior*, **25,** 131-147.

Mazur JE, Stellar JR and Waraczynski M (1987). Self-control choice with electrical stimulation of the brain as a reinforcer. *Behavioural Processes*, **15,** 143-153.

Mehlman PT, Higley JD, Faucher I*, et al.* (1994). Low CSF 5-HIAA concentrations and severe aggression and impaired impulse control in nonhuman primates. *American Journal of Psychiatry*, **151,** 1485-1491.

Mérö L (1998). *Moral Calculations: Game Theory, Logic, and Human Frailty*, Springer-Verlag, New York.

Mishkin M, Malamut B and Bachevalier J (1984). Memories and habits: two neural systems. In G Lynch, JL McGaugh and NM Weinberger, eds. *Neurobiology of Learning and Memory*, pp. 65-77. Guildford Press, New York.

Mitchell SH (1999). Measures of impulsivity in cigarette smokers and non-smokers. *Psychopharmacology*, **146,** 455-464.

Mobini S, Body S, Ho MY*, et al.* (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology*, **160,** 290-298.

Morrissey G, Ho MY, Wogar MA, Bradshaw CM and Szabadi E (1994). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on timing behaviour investigated with the fixed-interval peak procedure. *Psychopharmacology*, **114,** 463-468.

Morrissey G, Wogar MA, Bradshaw CM and Szabadi E (1993). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on timing behaviour investigated with an interval bisection task. *Psychopharmacology*, **112,** 80-85.

Mowrer OH (1960). *Learning Theory and Behavior*, Wiley, New York.

Muir JL, Everitt BJ and Robbins TW (1996). The cerebral cortex of the rat and visual attentional function: dissociable effects of mediofrontal, cingulate, anterior dorsolateral, and parietal cortex lesions on a five-choice serial reaction time task. *Cerebral Cortex*, **6,** 470-481.

Myerson J and Green L (1995). Discounting of delayed rewards: models of individual choice. *Journal of the Experimental Analysis of Behavior*, **64,** 263-276.

O'Brien CP, Childress AR, Ehrman R and Robbins SJ (1998). Conditioning factors in drug abuse: can they explain compulsion? *Journal of Psychopharmacology*, **12,** 15-22.

Packard MG and McGaugh JL (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, **65,** 65-72.

Papa M, Sagvolden T, Sergeant JA and Sadile AG (1996). Reduced CaMKII-positive neurones in the accumbens shell of an animal model of attention-deficit hyperactivity disorder. *Neuroreport*, **7,** 3017-3020.

Papa M, Sellitti S and Sadile AG (2000). Remodeling of neural networks in the anterior forebrain of an animal model of hyperactivity and attention deficits as monitored by molecular imaging probes. *Neuroscience and Biobehavioral Reviews*, **24,** 149-156.

Papa M, Sergeant JA and Sadile AG (1998). Reduced transduction mechanisms in the anterior accumbal interface of an animal model of attention-deficit hyperactivity disorder. *Behavioural Brain Research*, **94,** 187-195.

Parkinson JA, Cardinal RN and Everitt BJ (2000a). Limbic cortical-ventral striatal systems underlying appetitive conditioning. *Progress in Brain Research*, **126,** 263-285.

Parkinson JA, Robbins TW and Everitt BJ (1999). Selective excitotoxic lesions of the nucleus accumbens core and shell differentially affect aversive Pavlovian conditioning to discrete and contextual cues. *Psychobiology*, **27,** 256-266.

Parkinson JA, Willoughby PJ, Robbins TW and Everitt BJ (2000b). Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: Further evidence for limbic cortical-ventral striatopallidal systems. *Behavioral Neuroscience*, **114,** 42-63.

Pavlov IP (1927). *Conditioned Reflexes*, Oxford University Press, Oxford.

Paxinos G and Watson C (1998). *The Rat Brain in Stereotaxic Coordinates*. Fourth edition, Academic Press.

Pinker S (1997). *How The Mind Works*, Allen Lane - The Penguin Press, London.

Poulos CX, Le AD and Parker JL (1995). Impulsivity predicts individual susceptibility to high levels of alcohol self-administration. *Behavioural Pharmacology*, **6,** 810-814.

Poulos CX, Parker JL and Le AD (1996). Dexfenfluramine and 8-OH-DPAT modulate impulsivity in a delay-of-reward paradigm: implications for a correspondence with alcohol consumption. *Behavioural Pharmacology*, **7,** 395-399.

Rachlin H (1971). On the tautology of the matching law. *Journal of the Experimental Analysis of Behavior*, **15,** 249-251.

Rachlin H, Raineri A and Cross D (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, **55,** 233-244.

Revusky S and Garcia J (1970). Learned associations over long delays. In GH Bower, ed. *The Psychology of Learning and Motivation*, Vol. 4, pp. 1-84. Academic Press, New York.

Richards JB, Chock MA, Carlson B, de Wit H and Seiden L (1997a). Comparison of two models of impulsive behavior in rats: effects of amphetamine and haloperidol. *Society for Neuroscience Abstracts*, **23,** 2406.

Richards JB, Mitchell SH, de Wit H and Seiden LS (1997b). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior*, **67,** 353-366.

Richards JB, Sabol KE and de Wit H (1999). Effects of methamphetamine on the adjusting amount procedure, a model of impulsive behavior in rats. *Psychopharmacology*, **146,** 432-439.

Richards JB and Seiden LS (1995). Serotonin depletion increases impulsive behavior in rats. *Society for Neuroscience Abstracts*, **21,** 1693.

Richardson NR and Gratton A (1998). Changes in medial prefrontal cortical dopamine levels associated with response-contingent food reward: an electrochemical study in rat. *Journal of Neuroscience*, **18,** 9130-9138.

Robbins TW (1976). Relationship between reward-enhancing and stereotypical effects of psychomotor stimulant drugs. *Nature*, **264,** 57-59.

Robbins TW (1978). The acquisition of responding with conditioned reinforcement: effects of pipradrol, methylphenidate, d-amphetamine, and nomifensine. *Psychopharmacology*, **58,** 79-87.

Robbins TW and Everitt BJ (1992). Functions of dopamine in the dorsal and ventral striatum. *Seminars in the Neurosciences*, **4,** 119-127.

Robbins TW and Everitt BJ (1996). Neurobehavioural mechanisms of reward and motivation. *Current Opinion in Neurobiology*, **6,** 228-236.

Robbins TW, Watson BA, Gaskin M and Ennis C (1983). Contrasting interactions of pipradrol, d-amphetamine, cocaine, cocaine analogues, apomorphine and other drugs with conditioned reinforcement. *Psychopharmacology*, **80,** 113-119.

Roberts S (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, **7,** 242-268.

Robinson TE and Berridge KC (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews*, **18,** 247-291.

Rubia K, Overmeyer S, Taylor E*, et al.* (1999). Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: A study with functional MRI. *American Journal of Psychiatry*, **156,** 891-896.

Russell SJ and Norvig PN (1995). *Artificial Intelligence: a modern approach*, Prentice-Hall, Upper Saddle River, New Jersey.

Russell V, de Villiers A, Sagvolden T, Lamm M and Taljaard J (1998). Differences between electrically-, ritalin- and D-amphetamine- stimulated release of [H-3]dopamine from brain slices suggest impaired vesicular storage of dopamine in an animal model of attention-deficit hyperactivity disorder. *Behavioural Brain Research*, **94,** 163-171.

Russell V, Devilliers A, Sagvolden T, Lamm M and Taljaard J (1995). Altered dopaminergic function in the prefrontal cortex, nucleus accumbens and caudate-putamen of an animal model of attention-deficit hyperactivity disorder - the spontaneously hypertensive rat. *Brain Research*, **676,** 343-351.

Russell VA (2000). The nucleus accumbens motor-limbic interface of the spontaneously hypertensive rat as studied in vitro by the superfusion slice technique. *Neuroscience and Biobehavioral Reviews*, **24,** 133-136.

Sadile AG (2000). Multiple evidence of a segmental defect in the anterior forebrain of an animal model of hyperactivity and attention deficit. *Neuroscience and Biobehavioral Reviews*, **24,** 161-169.

Sagvolden T (2000). Behavioral validation of the spontaneously hypertensive rat (SHR) as an animal model of attention-deficit/hyperactivity disorder (AD/HD). *Neuroscience and Biobehavioral Reviews*, **24,** 31-39.

Sagvolden T, Aase H, Zeiner P and Berger D (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research*, **94,** 61-71.

Sagvolden T, Metzger MA, Schiorbeck HK, Rugland AL, Spinnangr I and Sagvolden G (1992). The spontaneously hypertensive rat (SHR) as an animal model of childhood hyperactivity (ADHD): changed reactivity to reinforcers and to psychomotor stimulants. *Behavioral and Neural Biology*, **58,** 103-112.

Sagvolden T, Pettersen MB and Larsen MC (1993). Spontaneously hypertensive rats (SHR) as a putative animal model of childhood hyperkinesis: SHR behavior compared to four other rat strains. *Physiology and Behavior*, **54,** 1047-1055.

Sagvolden T and Sergeant JA (1998). Attention deficit/hyperactivity disorder - from brain dysfunctions to behaviour. *Behavioural Brain Research*, **94,** 1-10.

Salamone JD (1994). The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. *Behavioural Brain Research*, **61,** 117-133.

Salamone JD, Cousins MS and Snyder BJ (1997). Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. *Neuroscience and Biobehavioral Reviews*, **21,** 341-359.

Schultz W, Apicella P, Romo R and Scarnati E (1995). Context-dependent activity in primate striatum reflecting past and future behavioral events. In JC Houk, JL Davis and DG Beiser, eds. *Models of Information Processing in the Basal Ganglia*, pp. 11-27. MIT Press, Cambridge, Massachusetts / London.

Schultz W, Apicella P, Scarnati E and Ljungberg T (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *Journal of Neuroscience*, **12,** 4595-4610.

Schultz W, Tremblay L and Hollerman JR (1998). Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*, **37,** 421-429.

Schultz W, Tremblay W and Hollerman JR (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex*, **10,** 272-283.

Shizgal P (1997). Neural basis of utility estimation. *Current Opinion in Neurobiology*, **7,** 198-208.

Solanto MV (1998). Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: a review and integration. *Behavioural Brain Research*, **94,** 127-152.

Sonuga-Barke EJS, Saxton T and Hall M (1998). The role of interval underestimation in hyperactive children's failure to suppress responses over time. *Behavioural Brain Research*, **94,** 45-50.

Soubrié P (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behavioral and Brain Sciences*, **9,** 319-335.

Spence KW (1956). *Behavior Theory and Conditioning*, Prentice-Hall, Englewood Cliffs, New Jersey.

Thorndike EL (1911). *Animal intelligence: experimental studies*, Macmillan, New York.

Tiffany ST and Drobes DJ (1990). Imagery and smoking urges: the manipulation of affective content. *Addictive Behaviors*, **15,** 531-539.

Tzschentke TM (2000). The medial prefrontal cortex as a part of the brain reward system. *Amino Acids*, **19,** 211-219.

von Neumann J and Morgenstern O (1947). *Theory of games and economic behavior*, Princeton University Press, Princeton, New Jersey.

Wade TR, de Wit H and Richards JB (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology*, **150,** 90-101.

White KG and Pipe ME (1987). Sensitivity to reinforcer duration in a self-control procedure. *Journal of the Experimental Analysis of Behavior*, **48,** 235-250.

White NM (1997). Mnemonic functions of the basal ganglia. *Current Opinion in Neurobiology*, **7,** 164-169.

Williams BA (1991). Marking and bridging versus conditioned reinforcement. *Animal Learning & Behavior*, **19,** 264-269.

Williams BA (1994a). Conditioned reinforcement: neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review*, **1,** 457-475.

Williams BA (1994b). Reinforcement and choice. In NJ Mackintosh, ed. *Animal Learning and Cognition*, pp. 81-108. Academic Press.

Williams BA and Dunn R (1991). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, **55,** 37-46.

Wise RA (1981). Brain dopamine and reward. In SJ Cooper, ed. *Theory in Psychopharmacology Volume 1*, pp. 103-122. Academic Press, London.

Wise RA (1982). Neuroleptics and operant behavior: the anhedonia hypothesis. *Behavioral and Brain Sciences*, **5,** 39-87.

Wise RA (1985). The anhedonia hypothesis: Mark III. *Behavioral and Brain Sciences*, **8,** 178-186.

Wise RA (1994). A brief history of the anhedonia hypothesis. In CR Legg and D Booth, eds. *Appetite: Neural and Behavioural Bases*, pp. 243-263. Oxford University Press, New York.

Wogar MA, Bradshaw CM and Szabadi E (1993a). Does the effect of central 5-hydroxytryptamine depletion on timing depend on motivational change? *Psychopharmacology*, **112,** 86-92.

Wogar MA, Bradshaw CM and Szabadi E (1993b). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on choice between delayed reinforcers. *Psychopharmacology*, **111,** 239-243.

Wultz B, Sagvolden T, Moser EI and Moser MB (1990). The spontaneously hypertensive rat as an animal model of attention-deficit hyperactivity disorder: effects of methylphenidate on exploratory behavior. *Behavioral and Neural Biology*, **53,** 88-102.

Wyvell CL and Berridge KC (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *Journal of Neuroscience*, **20,** 8122-8130.

Wyvell CL and Berridge KC (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered “wanting” for sucrose reward. *Journal of Neuroscience*, **21,** 7831-7840.

Figure 1

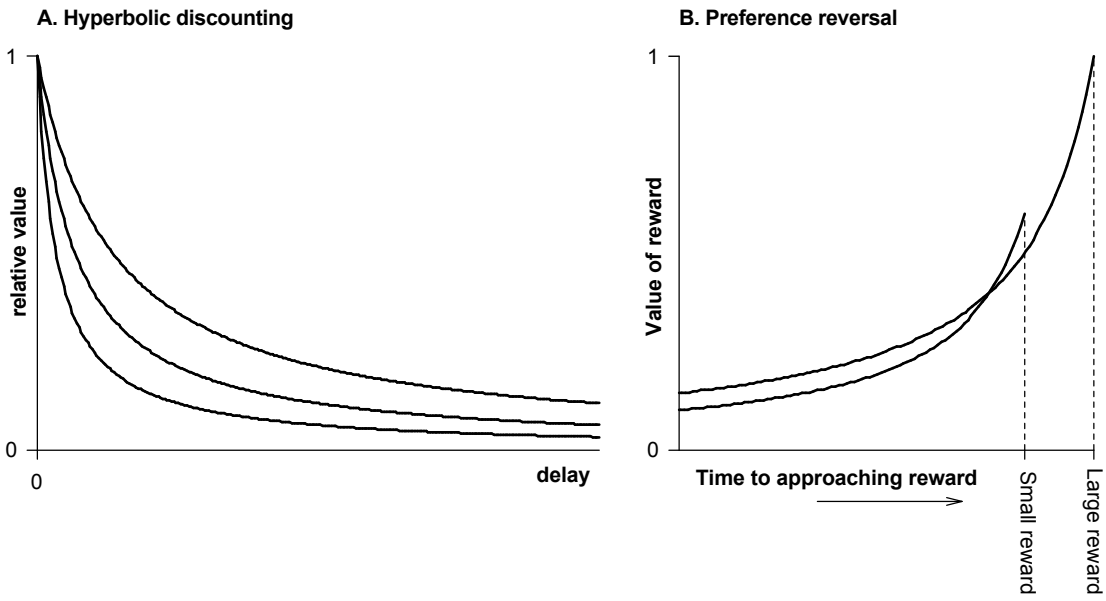
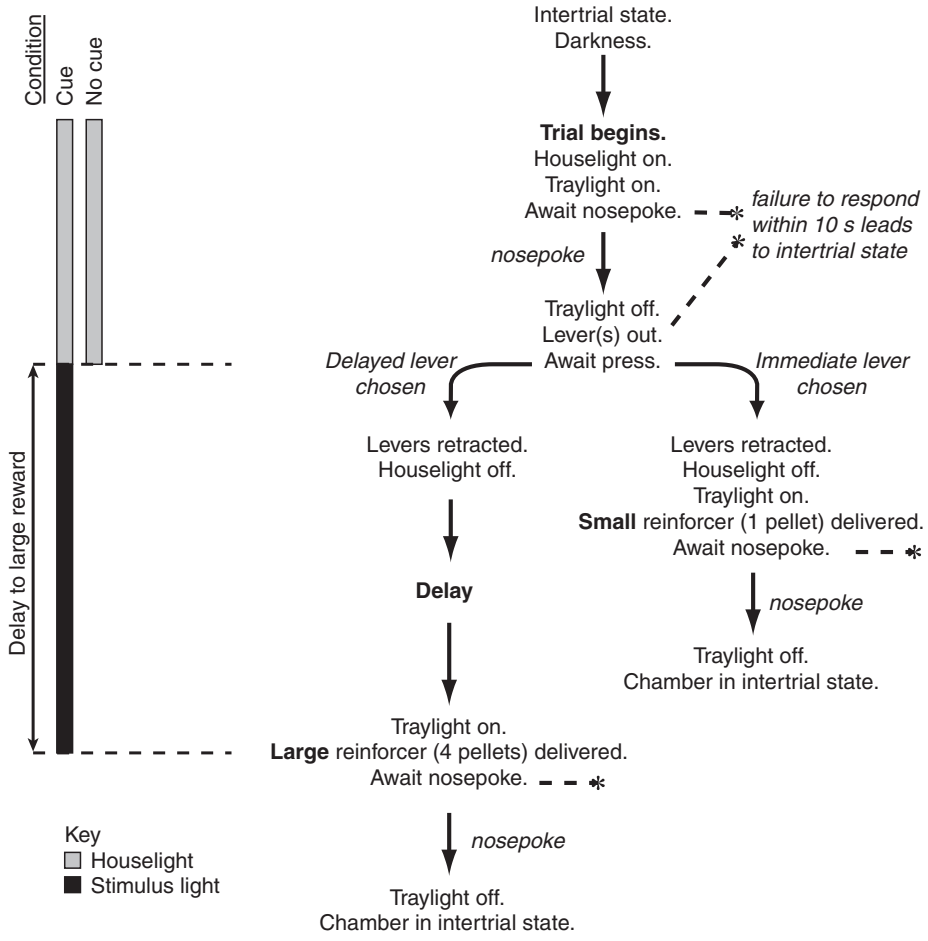

A. Hyperbolic discounting
1
relative value
0
0
delay
B. Preference reversal
1
Value of reward
0
Time to approaching reward
Small reward
Large reward

Figure 2


Condition
Cue
No cue
Delay to large reward
Key
Houselight
Stimulus light
Intertrial state.
Darkness.
Trial begins.
Houselight on.
Traylight on.
Await nosepoke.
failure to respond within 10 s leads to intertrial state
nosepoke
Traylight off.
Lever(s) out.
Await press.
Delayed lever chosen
Immediate lever chosen
Levers retracted.
Houselight off.
Delay
Traylight on.
Large reinforcer (4 pellets) delivered.
Await nosepoke.
nosepoke
Traylight off.
Chamber in intertrial state.
Levers retracted.
Houselight off.
Traylight on.
Small reinforcer (1 pellet) delivered.
Await nosepoke.
nosepoke
Traylight off.
Chamber in intertrial state.

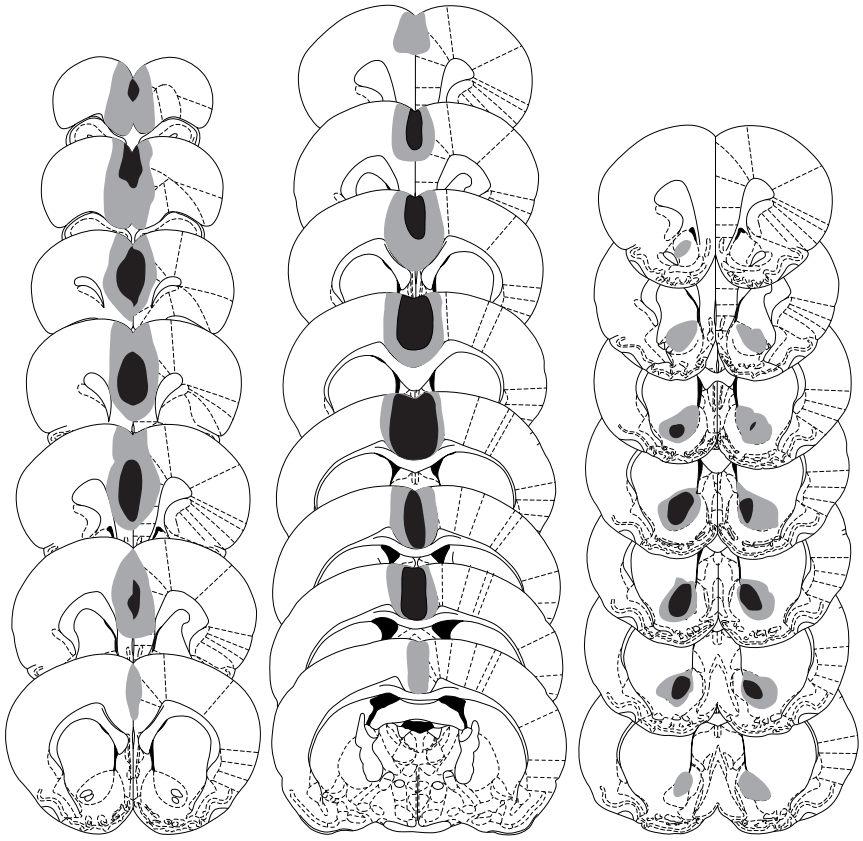


Figure 3

Figure 4

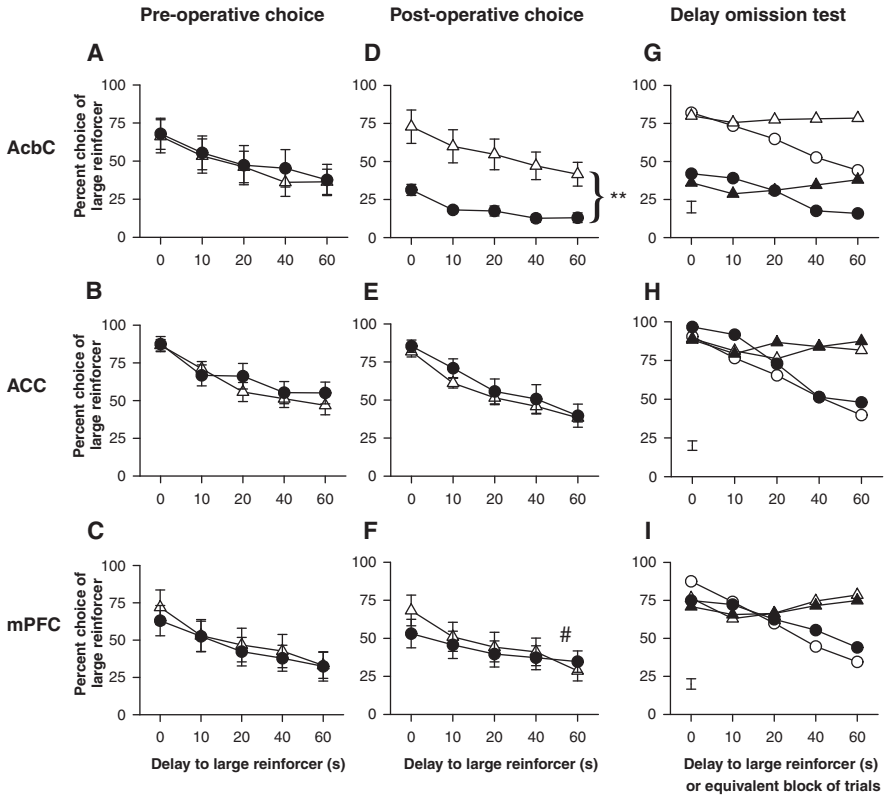
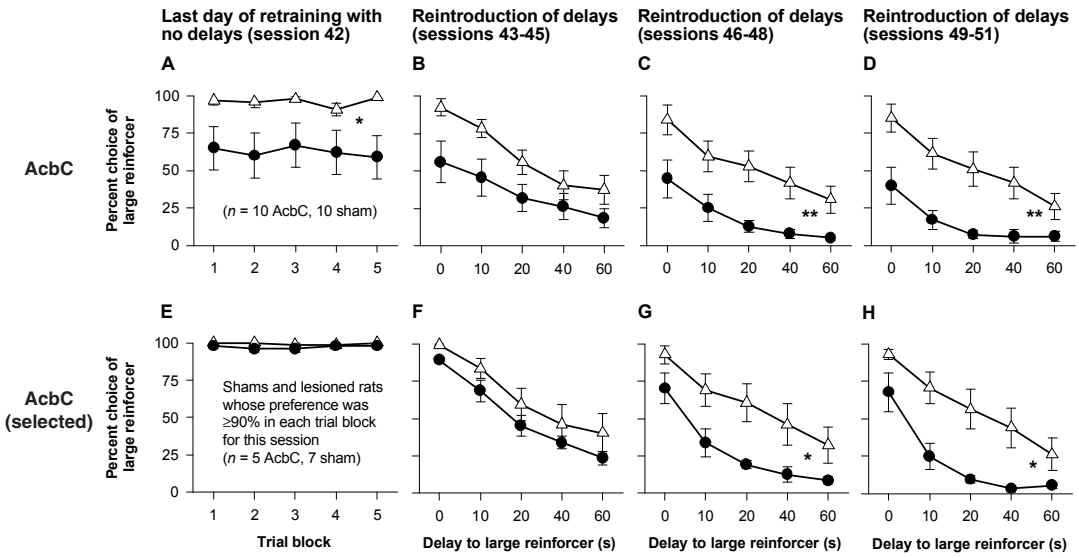

Pre-operative choice
Post-operative choice
Delay omission test
A
B
C
D
E
F
G
H
I
AcbC
ACC
mPFC
Percent choice of large reinforcer
100
75
50
25
0
0
10
20
40
60
**
#
Delay to large reinforcer (s)
Delay to large reinforcer (s)
Delay to large reinforcer (s) or equivalent block of trials

Figure 5


Last day of retraining with no delays (session 42)
Reintroduction of delays (sessions 43-45)
Reintroduction of delays (sessions 46-48)
Reintroduction of delays (sessions 49-51)
A
B
C
D
E
F
G
H
AcbC
AcbC (selected)
Percent choice of large reinforcer
100
75
50
25
0
*
**
(n = 10 AcbC, 10 sham)
Shams and lesioned rats whose preference was ≥90% in each trial block for this session (n = 5 AcbC, 7 sham)
1
2
3
4
5
Trial block
0
10
20
40
60
Delay to large reinforcer (s)