Neural systems involved in delay and risk assessment in the rat

A dissertation submitted for the degree of
Doctor of Medicine

Rudolf Nicholas Cardinal, MA PhD MB BChir
St John's College, Cambridge

May 2006

Department of Experimental Psychology
University of Cambridge
Downing Street
Cambridge CB2 3EB
United Kingdom

To Hannah

Contents

Preface

The following work was carried out at the Department of Experimental Psychology, University of Cambridge, during the years of 2002–2005.

I hereby declare that I have not submitted this dissertation, in whole or in part, for any other degree, diploma or qualification at any University. This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except where explicitly acknowledged. I have attempted to reference appropriately any idea or finding that is not my own.

Abstract

This thesis investigated the contribution of the nucleus accumbens core (AcbC) and the hippocampus (H) to choice and learning involving reinforcement that was delayed or unlikely. Animals must frequently act to influence the world even when the reinforcing outcomes of their actions are delayed. Learning with action–outcome delays is a complex problem, and little is known of the neural mechanisms that bridge such delays. Impulsive choice, one aspect of impulsivity, is characterized by an abnormally high preference for small, immediate rewards over larger delayed rewards, and is a feature of attention-deficit/hyperactivity disorder (ADHD), addiction, mania, and certain personality disorders. Furthermore, when animals choose between alternative courses of action, seeking to maximize the benefit obtained, they must also evaluate the likelihood of the available outcomes. Little is known of the neural basis of this process, or what might predispose individuals to be overly conservative or to take risks excessively (avoiding or preferring uncertainty, respectively), but risk taking is another aspect of the personality trait of impulsivity and is a feature of a number of psychiatric disorders, including pathological gambling and some personality disorders.

The AcbC, part of the ventral striatum, is required for normal preference for a large, delayed reward over a small, immediate reward (self-controlled choice) in rats, but the reason for this is unclear. Chapter 3 investigated the role of the AcbC in learning a free-operant instrumental response using delayed reinforcement, performance of a previously learned response for delayed reinforcement, and assessment of the relative magnitudes of two different rewards. Groups of rats with excitotoxic or sham lesions of the AcbC acquired an instrumental response with different delays (0, 10, or 20 s) between the lever-press response and reinforcer delivery. A second (inactive) lever was also present, but responding on it was never reinforced. The delays retarded learning in normal rats. AcbC lesions did not hinder learning in the absence of delays, but AcbC-lesioned rats were impaired in learning when there was a delay, relative to sham-operated controls. Rats were subsequently trained to discriminate reinforcers of different magnitudes. AcbC-lesioned rats were more sensitive to differences in reinforcer magnitude than sham-operated controls, suggesting that the deficit in self-controlled choice previously observed in such rats was a consequence of reduced preference for delayed rewards relative to immediate rewards, not of reduced preference for large rewards relative to small rewards. AcbC lesions also impaired the performance of a previously learned instrumental response in a delay-dependent fashion. These results demonstrate that the AcbC contributes to instrumental learning and performance by bridging delays between subjects' actions and the ensuing outcomes that reinforce behaviour.

When outcomes are delayed, they may be attributed to the action that caused them, or mistakenly attributed to other stimuli, such as the environmental context. Consequently, animals that are poor at forming context–outcome associations might learn action–outcome associations better with delayed reinforcement than normal animals. The hippocampus contributes to the representation of environmental context, being required for aspects of contextual conditioning. It was therefore hypothesized that animals with H lesions would be better than normal animals at learning to act on the basis of delayed reinforcement. Chapter 4 tested the ability of H-lesioned rats to learn a free-operant instrumental response using delayed reinforcement, and their ability to exhibit self-controlled choice. Rats with sham or excitotoxic H lesions acquired an instrumental response with different delays (0, 10, or 20 s) between the response and reinforcer delivery. H-lesioned rats responded slightly less than sham-operated controls in the absence of delays, but they became better at learning (relative to shams) as the delays increased; delays impaired learning less in H-lesioned rats than in shams. In contrast, lesioned rats exhibited impulsive choice, pre-

ferring an immediate, small reward to a delayed, larger reward, even though they preferred the large re-ward when it was not delayed. These results support the view that the H hinders action–outcome learning with delayed outcomes, perhaps because it promotes the formation of context–outcome associations in-stead. However, although lesioned rats were better at learning with delayed reinforcement, they were worse at choosing it, suggesting that self-controlled choice and learning with delayed reinforcement tax different psychological processes.

Chapter 5 examined the effects of excitotoxic lesions of the AcbC on probabilistic choice in rats. Rats chose between a single food pellet delivered with certainty (probability $p = 1$) and four food pellets deliv-ered with varying degrees of uncertainty ($p = 1$, 0.5, 0.25, 0.125, and 0.0625) in a discrete-trial task, with the large-reinforcer probability decreasing or increasing across the session. Subjects were trained on this task and then received excitotoxic or sham lesions of the AcbC before being retested. After a transient period during which AcbC-lesioned rats exhibited relative indifference between the two alternatives com-pared to controls, AcbC-lesioned rats came to exhibit risk-averse choice, choosing the large reinforcer less often than controls when it was uncertain, to the extent that they obtained less food as a result. Rats behaved as if indifferent between a single certain pellet and four pellets at $p = 0.32$ (sham-operated) or at $p = 0.70$ (AcbC-lesioned) by the end of testing. When the probabilities did not vary across the session, AcbC-lesioned rats and controls strongly preferred the large reinforcer when it was certain, and strongly preferred the small reinforcer when the large reinforcer was very unlikely ($p = 0.0625$), with no differ-ences between AcbC-lesioned and sham-operated groups. These results suggest that the AcbC contributes to action selection by promoting the choice of uncertain, as well as delayed, reward.

Key words:

  delay
  uncertainty
  impulsivity
  addiction
  nucleus accumbens
  hippocampus

Acknowledgements

Abbreviations

$\tilde{\varepsilon}$	Huynh–Feldt epsilon
ε	price elasticity
(a, b)	a range a–b that includes neither a nor b, i.e. a range $a < x < b$.
$[a, b)$	a range a–b that includes a but not b, i.e. a range $a \leq x < b$.
$[a, b]$	a range a–b that includes both a and b, i.e. a range $a \leq x \leq b$.
5-HIAA	5-hydroxyindoleacetic acid
5-HT	5-hydroxytryptamine (serotonin)
Acb	nucleus accumbens
AcbC	nucleus accumbens core
AcbSh	nucleus accumbens shell
ADHD	attention-deficit/hyperactivity disorder
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolpropionate
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AP-5	D-(−)-2-amino-5-phosphonopentanoic acid
BLA	basolateral amygdala
BOLD	blood oxygen level dependent (of an fMRI signal)
CA	cornu ammonis (Ammon's horn)
cf.	*confer* (compare)
ch.	chapter
COD	changeover delay
CPP	3-(2-carboxypiperazin-4-yl)-propyl-1-phosphonic acid
CR	conditioned response
CRH	corticotrophin-releasing hormone (also known as corticotrophin-releasing factor, CRF)
CS	conditioned stimulus
CSF	cerebrospinal fluid
DA	dopamine
df	degrees of freedom
DRL	differential reinforcement of low rates
DRO	differential reinforcement of other behaviour
ECS	electroconvulsive shock (synonym for ECT)
ECT	electroconvulsive therapy (synonym for ECS)
e.g.	*exempli gratia* (for example)
et al.	and others (*et alii*, masculine plural; *et aliae*, feminine plural; *et alia*, neutral plural)
etc.	*et cetera* (and the rest)
FI	fixed interval
fMRI	functional magnetic resonance imaging
FR	fixed ratio
h	hour
H	hippocampus
i.e.	*id est* (that is to say)
i.m.	intramuscular

i.p.	intraperitoneal
ICSS	intracranial self-stimulation
ISI	interstimulus interval
ITI	intertrial interval
L.	Latin for
LL	larger, later (in the context of rewards)
LTP	long-term potentiation
LTD	long-term depression
min	minute
mPFC	medial prefrontal cortex
MRI	magnetic resonance imaging
n	number of subjects or observations
NA	noradrenaline
NMDA	*N*-methyl-D-aspartate
NS	not significant
OCD	obsessive–compulsive disorder
OFC	orbitofrontal cortex
p	probability
P(A)	probability of event A occurring
P(A│B)	probability of A occurring, given that B has occurred
p., pp.	page, pages
PBS	phosphate-buffered saline
PFC	prefrontal cortex
PIT	Pavlovian–instrumental transfer
PKA	protein kinase A (cyclic-adenosine-monophosphate-dependent protein kinase)
$p_{reinforcer}$	probability of delivery of a reinforcer after it has been chosen
$p_{statistical}$	statistical *p* value (probability of obtaining the observed data, or results more extreme, were the null hypothesis to be true)
q.v.	*quod vide* (which see)
r^2	proportion of variance explained
RI	random interval
RR	random ratio
SED	standard error of the difference between means
SEM	standard error of the mean
SHR	spontaneously hypertensive rat
SNc	substantia nigra pars compacta
S–R	stimulus–response
SS	sum of squares (sum of squared deviations from a mean) (in the context of statistics)
SS	smaller, sooner (in the context of rewards)
STN	subthalamic nucleus
TCP/IP	transmission control protocol/internet protocol
US	unconditioned stimulus
v.	versus

v/v	volume per unit volume[1]
VR	variable ratio
VTA	ventral tegmental area
w/v	weight per unit volume

---

[1] Concentrations given as percentages are calculated as follows. A 1% solution, volume per unit volume (v/v), is a solution in which $^1/_{100}$ of the total volume is solute. A 1% solution, weight by unit weight (w/w), is one in which 1% of the total weight of the solution is solute; thus, a 1% solution implies 1 g of solute dissolved in 99 g of solvent. A 1% solution, weight by unit volume (w/v), is a solution of 1 g in a total volume of 100 ml (10 g l^{-1}); "100%" denotes 1 kg l^{-1}. Similarly, the notation "1:1000" denotes 1 g l^{-1} (1 mg ml^{-1}).

Publications

The publications listed below are submitted in support of this dissertation, under Regulation 7 of the Ordinances of the University of Cambridge concerning the degree of Doctor of Medicine (Ordinances, Chapter 7, at http://www.admin.cam.ac.uk/univ/so/so_ch07.pdf). These publications do not form part of work I have submitted for any other degree, diploma or qualification at any University. Those marked (*) are central to the material presented in this thesis.

Articles indexed by digital object identifier (DOI) can be retrieved electronically from the publisher: if the DOI is xxx, the URL is <http://dx.doi.org/xxx>. An up-to-date publication list, with electronic copies, is at <http://pobox.com/~rudolf/publications>.

2001

1. **Rahman S, Sahakian BJ, Cardinal RN, Rogers RD, Robbins TW** (2001). Decision-making and neuropsychiatry. *Trends in Cognitive Sciences* **5:** 271–277. DOI 10.1016/S1364-6613(00)01650-8.

2. **Dalley JW, McGaughy J, O'Connell MT, Cardinal RN, Levita L, Robbins TW** (2001). Distinct changes in cortical acetylcholine and noradrenaline efflux during contingent and non-contingent performance of a visual attentional task. *Journal of Neuroscience* **21:** 4908–4914.

3. **Di Ciano P, Cardinal RN, Cowell RA, Little SJ, Everitt BJ** (2001). Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of Pavlovian approach. *Journal of Neuroscience* **21:** 9471–9477.

4. **Cardinal RN, Aitken MRF** (2001). *Whisker*, version 2, computer software. Described in detail at <http://www.whiskercontrol.com>.

2002

5. **Cardinal RN, Parkinson JA, Lachenal G, Halkerston KM, Rudarakanchana N, Hall J, Morrison CH, Howes SR, Robbins TW, Everitt BJ** (2002). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience* **116:** 553–567. DOI 10.1037//0735-7044.116.4.553.

6. **Parkinson JA, Dalley JW, Cardinal RN, Bamford A, Fenhert B, Lachenal G, Rudarakanchana N, Halkerston KM, Robbins TW, Everitt BJ** (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behavioural Brain Research* **137:** 149–163. DOI 10.1016/S0166-4328(02)00291-7.

2003

7. **Cardinal RN, Robbins TW, Everitt BJ** (2003). Choosing delayed rewards: perspectives from learning theory, neurochemistry, and neuroanatomy. Chapter 6 (and reply to commentary) of Vuchinich RE, Heather N (eds), *Choice, Behavioral Economics and Addiction* (proceedings of a conference in Birmingham, Alabama, USA; 15–17 March 2002). Elsevier, Amsterdam (ISBN 0080440568), pp. 183–213 and 217–218. (*)

8. **Everitt BJ, Cardinal RN, Parkinson JA, Robbins TW** (2003). Appetitive behavior: the impact of amygdala-dependent mechanisms of emotional learning. *Annals of the New York Academy of Sciences* **985:** 233–250.

2004

9. **Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ** (2004). Limbic corticostriatal systems and delayed reinforcement. *Annals of the New York Academy of Sciences* **1021:** 33–50. DOI `10.1196/annals.1308.004.`(*)

10. **Cardinal RN, Everitt BJ** (2004). Neural and psychological mechanisms underlying appetitive learning: links to drug addiction. *Current Opinion in Neurobiology* **14**(2): 156–162. DOI `10.1016/j.conb.2004.03.004.`

11. **Dalley JW, Cardinal RN, Robbins TW** (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neuroscience and Biobehavioral Reviews* **28:** 771–784. DOI `10.1016/j.neubiorev.2004.09.006.`

12. **Cardinal RN** (2004). Waiting for better things. *The Psychologist* **17:** 684–687.

13. **Winstanley CA, Theobald DEH, Cardinal RN, Robbins TW** (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *Journal of Neuroscience* **24:** 4718–4722. DOI `10.1523/JNEUROSCI.5606-03.2004.`

14. **Dalley JW, Theobald DE, Bouger P, Chudasama Y, Cardinal RN, Robbins TW** (2004). Cortical cholinergic function and deficits in visual attentional performance in rats following 192 IgG-saporin-induced lesions of the medial prefrontal cortex. *Cerebral Cortex* **14:** 922–932. DOI `10.1093/cercor/bhh052.`

2005

15. **Cardinal RN, Cheung THC** (2005). Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat. *BMC Neuroscience* **6:** 9. DOI `10.1186/1471-2202-6-9.`(*)

16. **Winstanley CA, Theobald DEH, Dalley JW, Cardinal RN, Robbins TW** (2005). Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cerebral Cortex* **16:** 106–114 (advance access 13 April 2005). DOI `10.1093/cercor/bhi088.`

17. **Cheung THC, Cardinal RN** (2005). Hippocampal lesions facilitate instrumental learning with delayed reinforcement but induce impulsive choice in rats. *BMC Neuroscience* **6:** 36. DOI `10.1186/1471-2202-6-36.`(*)

18. **Cardinal RN, Howes NJ** (2005). Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. *BMC Neuroscience* **6:** 37. DOI `10.1186/1471-2202-6-37.`(*)

19. **Robbins TW, Cardinal RN, Di Ciano P, Halligan PWG, Hellemans KGC, Lee JCL, Everitt BJ** (2005). Neuroscience of drugs and addiction. UK Office of Science and Technology *Foresight: Brain Science, Addiction and Drugs* project, 13 July 2005; <http://www.foresight.gov.uk>. [The OST is part of the Department of Trade and Industry; the project was also supported by the Home Office and the Department of Health.] (*)

2006

20. **Cardinal RN, Aitken MRF** (2006). *ANOVA for the behavioural sciences researcher.* Lawrence Erlbaum Associates, New Jersey (ISBN 0805855858 hardback or 0805855866 paperback). Supporting Web site at <http://www.psychol.cam.ac.uk/statistics>.

21. **Cardinal RN** (2006, in press). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Networks*. (*)

List of figures

List of tables

Chapter 1: Introduction

1.1 OVERVIEW

The problems of learning with delayed and uncertain reinforcement, and of choosing delayed or uncertain rewards, are interesting from a theoretical and a practical perspective.

The theoretical perspective concerns the neural mechanisms of learning and choice. Animals act to obtain rewards including food, shelter, and sex. Sometimes, their actions are rewarded or reinforced immediately, but often this is not the case. Often, natural reinforcers follow the action that obtains them by a delay, even if it is short; to be successful, animals must learn to bridge these delays and act on the basis of delayed reinforcement. They may also profit by choosing delayed reinforcers over more immediate reinforcers, if the delayed reinforcers are sufficiently large. Likewise, animals that can act despite uncertainty as to what the future holds and can calculate risk appropriately are placed at a competitive advantage.

The practical perspective concerns individual variation in sensitivity to delayed and uncertain reinforcement. Individuals differ in their ability to choose delayed rewards. Self-controlled individuals are strongly influenced by delayed reinforcement, and choose large, delayed rewards in preference to small, immediate rewards; in contrast, individuals who are relatively insensitive to delayed reinforcement choose impulsively, preferring the immediate, smaller reward in this situation (Ainslie, 1975). Impulsivity has long been recognized as a normal human characteristic (Aristotle, 350 BC / 1925) and in some circumstances it may be beneficial (Evenden, 1999b), 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 attention-deficit/hyperactivity disorder (ADHD) (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998). Similar individual differences in the tendency to work for uncertain rewards may contribute to personality traits such as venturesomeness (Eysenck & Eysenck, 1977; Evenden, 1999a), but risk taking is another aspect of impulsivity (Daruna & Barnes, 1993; Eysenck, 1993; Evenden, 1999a) and is a feature of a number of psychiatric disorders, including pathological gambling and certain personality disorders (Roy *et al.*, 1989; Coccaro & Siever, 1995; APA, 2000; Holt *et al.*, 2003).

In this chapter, top-down (behavioural economic) and bottom-up (animal learning theory) approaches to action and decision making are outlined. The specific problems that delays and uncertainty cause for learning and choice are reviewed, with a discussion of the relevance of individual differences in delay and uncertainty processing. The effects of delays and uncertainty upon learning and choice in normal animals are reviewed. Next, systemic psychopharmacological studies examining delayed or uncertain reinforcement are discussed. Neurobiological studies are then examined, beginning with a brief review of the anatomy of relevant parts of the limbic system, followed by an overview of the role of these structures in reinforcement learning in general, and concluding with a review of neuroanatomical studies specifically concerning delayed and/or uncertain reinforcement. An overview of experimental work contained in this thesis is then provided.

1.2 NORMATIVE AND BEHAVIOURAL ECONOMIC APPROACHES TO DECISION MAKING

Behavioural economics is a merging of traditional economic theory with psychological studies of choice (Rachlin *et al.*, 1976; Allison, 1979) that offers a quantitative approach to choice and decision making.

Much of economics is based on utility theory (von Neumann & Morgenstern, 1947; Russell & Norvig, 1995), which assumes that agents are rational in that they exhibit certain reasonable attributes of preference. For example, one assumption is transitivity of preference: if an agent prefers A to B and B to C, then it must prefer A to C (or it would easily be exploited by more rational agents). Given these assumptions, there must exist a utility function that assigns unidimensional values to real-world multidimensional events or outcomes, such that the agent prefers outcomes with higher utility. Psychologically and neurally, a similar process must also happen (Shizgal, 1997)—if at no earlier stage of processing, incompatible behaviours must compete for access to motor output. Agents can then use their knowledge about the world, and about the consequences of their actions (which may be uncertain), to act so as to maximize their expected utility (Arnauld & Nicole, 1662). 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 $P(action \rightarrow outcome_n | evidence)$ denoting the probability, given the available evidence, that *action* causes $outcome_n$. If $U(outcome_n)$ is the utility of obtaining $outcome_n$, then the expected utility of an action is therefore given by $EU(action | evidence) = \sum_n P(action \rightarrow outcome_n | evidence) \cdot U(outcome_n)$. 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. Rational behaviour need not require complex, explicit thought; merely that observed behaviour follows rational principles.

Conversely, if agents are logical, then we can infer their value system (utility function) by observing their behaviour—the principle of revealed preference (Friedman, 1990; Williams, 1994). To do so, we must assume that agents have reasonably simple objectives (for if we allow that the behaviour we observe is itself the agent's objective, we could explain any arbitrary behaviour).

Assuming rationality allows us to predict behaviour much better than not assuming rationality, unless we can predict the specific way in which people will be irrational (Friedman, 1990). However, humans do not always choose according to rational norms. Introducing an element of randomness into decision making can be theoretically optimal in some situations (von Neumann & Morgenstern, 1947; Mérö, 1998) and the requirement to make rapid decisions may promote the use of heuristics to approximate rational decision making (Russell & Norvig, 1995). Empirically, humans systematically deviate from the optimum when making decisions (Kahneman *et al.*, 1982; Heckerman *et al.*, 1992; Lopes, 1994; Chase *et al.*, 1998; Mullainathan, 2002). They do so because human cognitive abilities are limited ("bounded rationality") and because people frequently make choices that aren't in their long-term interest ("bounded willpower"). In particular, humans and animals do not discount the future in a self-consistent way (Ainslie, 1975; 2001). This point will be expanded upon later, but it serves to illustrate the departure of animal decision making from economic optimality in some situations.

1.3 BASIC PSYCHOLOGY OF REINFORCEMENT LEARNING

An alternative perspective on actions and their consequences stems from the animal learning theory literature. Whereas behavioural economists tend to adopt a top-down approach, treating the agent as a single decision-making entity, animal learning theorists have sought to identify subcomponents and mechanisms giving rise to overt behaviour. The study of motivated action is the study of instrumental conditioning—the process by which animals alter their behaviour when there is a contingency between their behaviour and a reinforcing outcome (Thorndike, 1911). Reinforcement learning (Minsky, 1961; Russell & Norvig, 1995; Haykin, 1999) has been studied for a long time (Thorndike, 1905; Thorndike, 1911; Grindley,

1932; Guthrie, 1935; Skinner, 1938; Hull, 1943). At its most basic level, it is the ability to learn to act on the basis of important outcomes such as reward and punishment; events that strengthen (increase the likelihood of) preceding responses are called positive reinforcers, and events whose removal strengthens preceding responses are called negative reinforcers (Skinner, 1938; 1953). If reinforcers are defined by their effect on behaviour, then, to avoid a circular argument, behaviour cannot be said to have altered as a consequence of reinforcement (Skinner, 1953). However, to explain behaviour, rather than merely to describe it, internal processes such as motivation must also be accounted for. Central motivational states, such as hunger and thirst, are intervening variables that parsimoniously account for a great deal of behavioural variability (Erwin & Ferguson, 1979; Toates, 1986; Ferguson, 2000). For example, water deprivation, eating dry food, hypertonic saline injection, and the hormone angiotensin II all induce a common state (thirst) that has multiple effects: thirsty animals drink more water, drink water faster, perform more of an arbitrary response to gain water, and so on. The ideas of motivational state entered early theories of reinforcement. For example, it was suggested that events that reduce "drive" states such as thirst are positively reinforcing (Hull, 1943). However, on its own this simple model cannot account for many instrumental conditioning phenomena, let alone "unnatural" reinforcement such as intracranial self-stimulation (ICSS) and drug addiction.

Modern neuropsychological theories of instrumental conditioning recognize that many processes contribute to a simple act such as pressing a lever to receive food (e.g. Dickinson, 1994). I will merely summarize these processes here; for a full review, see Cardinal *et al.* (2002a). Rats and humans exhibit goal-directed action, which is based on knowledge of the contingency between one's actions and their outcomes, and knowledge of the value of those outcomes. These two knowledge representations interact so that we work for that which we value (Dickinson, 1994; Dickinson & Balleine, 1994). Environmental stimuli ("discriminative stimuli" or S^{D}s) provide information about what contingencies may be in force in a given environment (Colwill & Rescorla, 1990; Rescorla, 1990a; 1990b). Remarkably, the value system governing goal-directed action is not the brain's only one. This "cognitive" value system (sometimes termed "instrumental incentive value") may be distinguished and dissociated (Balleine & Dickinson, 1991) from a different valuation process that determines our reactions when we actually experience a goal such as food—termed "liking", "hedonic reactions", or simply "pleasure" (Garcia, 1989). Under many normal circumstances the two values reflect each other and change together. However, the fact that they are different means that animals must *learn* what outcomes are valuable (hedonically pleasant) in a given motivational state, a process referred to as incentive learning. For example, rats do not know that to eat a particular food while sated is not as valuable as to eat the same food while hungry *until* they have actually eaten the food while sated (Balleine, 1992).

Just as there is more than one value system, there is more than one route to action. Not all action is goal directed. With time and training, actions can become habitual (Adams, 1982)—that is, elicited in relevant situations by direct stimulus–response (S–R) associations. S–R habits are less flexible than goal-directed action, because their representation contains no information about what the outcome will be, and therefore cannot alter quickly if the desirability of a particular outcome changes. However, habits may be important to reduce the demands on the cognitive, goal-directed system in familiar settings.

Finally, environmental stimuli have effects beyond eliciting habits and serving as discriminative stimuli. Stimuli that predict reward may become conditioned stimuli (CSs), associated with the reward (unconditioned stimulus, US) through Pavlovian associative learning (Pavlov, 1927). Pavlovian CSs can elicit Pavlovian conditioned responses (CRs), can influence ongoing instrumental behaviour directly (termed Pavlovian–instrumental transfer or PIT), and can serve as the goals of behaviour (termed condi-

tioned reinforcement) (see Estes, 1948; Lovibond, 1983; Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a).

1.4 DELAYED AND UNCERTAIN REINFORCEMENT: THE PROBLEMS OF LEARNING AND CHOICE

Natural and artificial learning agents must grapple with the problem of selecting actions to achieve the best possible outcome under their value system. However, the outcome of a given action is not always certain and immediate. Outcomes are frequently uncertain: agents do not always obtain that for which they work. Furthermore, when an agent acts to obtain reward or reinforcement, there is often a delay between its action and the ultimate outcome. This applies both to positive reinforcers (things whose presentation increases the likelihood of preceding actions) and negative reinforcers (things whose removal increases the likelihood of preceding actions) (Skinner, 1938), though I will focus on positive, or appetitive, reinforcers, such as food; I will also use the term "reward" for an appetitive positive reinforcer. For optimal performance, therefore, agents must learn and choose on the basis of reward or reinforcement that is uncertain or delayed.

As discussed above, agents may act procedurally, meaning that they act without an representation of the outcome of their actions, but merely on the basis that an action has been reinforced or led to unspecified "good things" before. Direct links between representations of triggering stimuli and particular responses exemplify procedural responding, or stimulus–response (S–R) learning; the S–R links are strengthened in some way as a result of the arrival of reinforcement, but without the nature of that reinforcement being explicitly encoded. Alternatively, or additionally, agents may encode the outcomes of their actions explicitly, and use these explicit (sometimes termed declarative) representations of anticipated actions when choosing what to do. Animals exhibit both stimulus–response (procedural) and truly goal-directed or action–outcome (declarative) responding (Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a). This complicates the analysis of motivated behaviour in animals, including the analysis of learning with and choosing uncertain and delayed rewards.

In an S–R learning system, it is easy to envisage connectionist mechanisms by which uncertain and delayed reinforcers could drive learning. Suppose an agent experiences its world, causing many different "stimulus units" to become activated, and suppose it acts randomly by activating different "response units". Let us consider the basic case of appetitive, certain, immediate reinforcement. Suppose a hardwired mechanism exists to detect events of innate importance to the agent, such as food to a hungry animal. Suppose also that this mechanism, upon detecting an important appetitive event, triggers an internal reinforcement signal that acts to strengthen links between currently active units (stimulus units and response units). By strengthening links between units representing the stimuli currently being perceived and the response currently executing, this simple system would reinforce the response, i.e. increase the probability of executing the response again in the same situation. These S–R links do not encode the nature of the food. If the relationship between responses and food is uncertain, i.e. if $0 < P(outcome \mid action) < 1$, then S–R connections will be reinforced on occasions when food is delivered, but not reinforced on occasions when it is not. S–R links would thus develop to reflect the statistical relationships between actions and reward in a particular stimulus environment: more reliable action–outcome contingencies in the environment come to be reflected in stronger S–R links. To extend this to delayed reinforcement, when the time $t(action \rightarrow outcome) > 0$, requires that some representation of recently executed responses remains active until the reinforcing outcome actually arrives, if the correct response is to be reinforced. If the ac-

tion representation decays gradually (or if it persists until a new action is begun, and the probability of remaining in the same "action state" therefore declines with time), then the likelihood of reinforcing the correct response will decline gradually as action–outcome delays increase, and the agent will learn less well as reinforcement is progressively delayed. None of these ideas are new (Thorndike, 1911; Grindley, 1932; Hull, 1932; Guthrie, 1935; Hull, 1943; Spence, 1956; Mowrer, 1960; Revusky & Garcia, 1970; Mackintosh, 1974; Killeen & Fetterman, 1988).

In a goal-directed (action–outcome) learning system, the agent must encode both the action–outcome relationship and the value of the outcome, and these two representations must interact to determine the probability of selecting a given action (Tolman, 1932; Dickinson, 1980; Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a). Declarative representations are substantially harder to represent using a simple connectionist framework (Holyoak & Spellman, 1993; Shastri & Ajjanagadde, 1993; Sougné, 1998). The problem of detecting and encoding the action–outcome relationship (the consequences of the agent's actions) is itself complex, but the additional issues concerning uncertain and delayed outcomes are much the same as in the S–R case. That is to say, it may be more difficult to learn that an action causes a given outcome if that outcome is inconsistent or delayed. On top of this, even if the agent knows perfectly well that an action produces an outcome with a certain probability and/or a certain delay, the agent may *value* uncertain or delayed rewards less than certain or immediate rewards, reducing the likelihood of its choosing that action. For example, if we ask a man whether he prefers £10 now or £20 next week, we usually assume that he represents the action–outcome contingencies equally (i.e. that he believes that selecting the "£10 now" option is as likely to produce £10 now as selecting the "£20 next week" option is to produce £20 next week) and that his choice simply reflects the relative value to him of the two options. On the other hand, if we train rats to press levers for (say) immediate and delayed reward, we must bear in mind the possibility of inequalities both in the representation of the action–outcome contingency for immediate and delayed reward, and in the values of the two outcomes—not to mention differences in S–R learning that the delays may engender. There are few mechanistic models of explicit (declarative) delay or uncertainty coding applicable to animal learning, although a recent model proposes the encoding of action uncertainty as a way of mediating competition between goal-directed and S–R responding (Daw *et al.*, 2005).

1.5 INDIVIDUAL DIFFERENCES: RISK TAKING AND IMPULSIVITY

Individual differences in responsivity to uncertain or delayed reinforcement are also of considerable interest. When making decisions under conditions of uncertainty, individuals vary as to how much uncertainty or risk they are willing to tolerate. Formally, individuals differ in how much they discount the value of reinforcers as the uncertainty of the reinforcer increases—i.e. as the probability of the reinforcer declines, or the odds against obtaining the reinforcer increase (Ho *et al.*, 1999). Risk taking is one aspect of the personality trait of impulsivity (Daruna & Barnes, 1993; Eysenck, 1993; Evenden, 1999a) and is a feature of a number of psychiatric disorders, including pathological gambling, antisocial personality disorder, and borderline personality disorder (Roy *et al.*, 1989; Coccaro & Siever, 1995; APA, 2000; Holt *et al.*, 2003). The term "risk" implies exposure to the possibility of an aversive consequence (OUP, 1997), which may include the possibility of not obtaining an anticipated reward. In the appetitive domain, risk taking is exemplified by the tendency to choose large rewards that are very uncertain, in preference to smaller, certain rewards. Abnormal risk taking may reflect dysfunction of reinforcement learning systems that mediate the effects of uncertain reward or punishment.

Furthermore, individual variation in the ability to use delayed reinforcement may determine another aspect of impulsivity: an animal able to forgo short-term poor rewards in order to obtain delayed but better rewards may be termed self-controlled, whereas an animal that cannot tolerate delays to reward may be said to exhibit impulsive choice (Ainslie, 1975; Evenden, 1999b; Evenden, 1999a; Ainslie, 2001). Abnormalities in learning from delayed reinforcement may be of considerable clinical significance (Rahman *et al.*, 2001). Impulsivity is part of the syndrome of many psychiatric disorders, including mania, drug addiction, antisocial personality disorder, borderline personality disorder, and ADHD (APA, 2000). Impulsivity is a broad concept that may be divided into preparation impulsivity (failure to take all relevant information into account before making a decision), execution or "motor" impulsivity (termination of a behavioural chain before the goal is reached), and outcome or "choice" impulsivity (choice of a quick but less valuable outcome rather than a later but more valuable outcome). These measures may be pharmacologically dissociated (Evenden, 1999b; Evenden, 1999a). Impulsive choice, one aspect of impulsivity (Evenden, 1999b), may reflect dysfunction of reinforcement learning systems mediating the effects of delayed rewards (Ainslie, 1975; Sagvolden & Sergeant, 1998).

1.6 LEARNING WITH DELAYED REINFORCEMENT IN NORMAL ANIMALS

1.6.1 Basic phenomena

Delays can hamper both Pavlovian and instrumental conditioning (Dickinson, 1980; Mackintosh, 1983; Dickinson, 1994; Gallistel, 1994; Hall, 1994). For example, instrumental conditioning has long been ob-
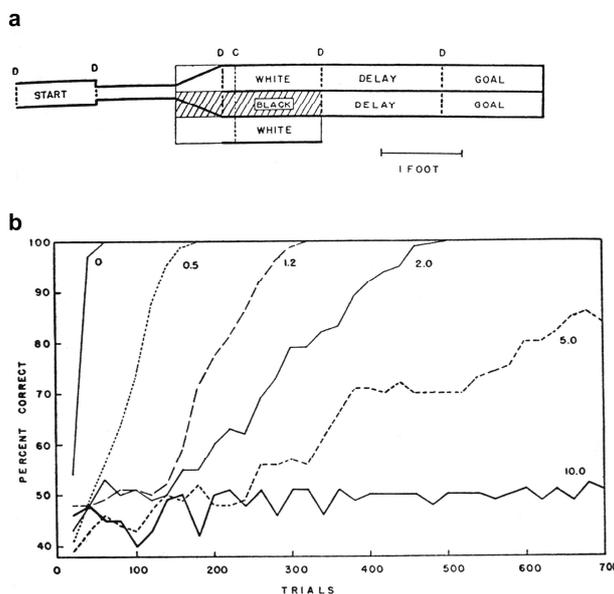
FIG. 2. Learning curves for each of the six different delay groups
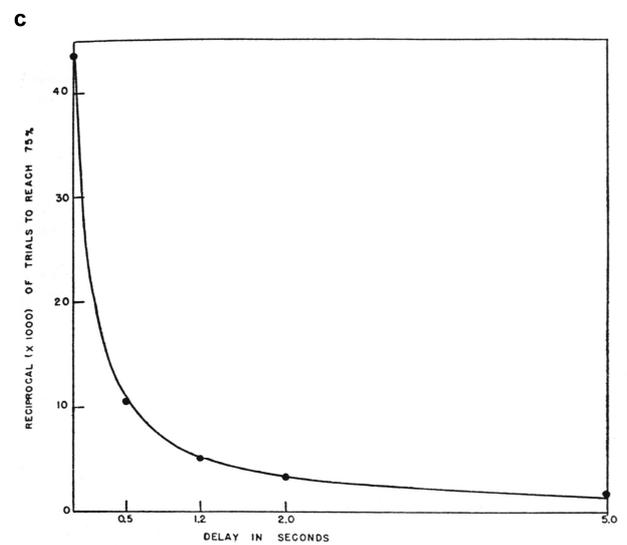
FIG. 3. Rate of learning as a function of delay of reward. The reciprocal × 1000 of the number of trials to reach the level of 75 percent correct choices is plotted against the time of delay. Experimental values are represented by black dots and the smooth curve is fitted to these data.

Figure 1: Discrimination learning with delayed reinforcement
Grice (1948) trained rats on a visual discrimination task with delayed reinforcement. The rats had a choice of a white or a black start alley (which varied in their left/right position); the delay was provided by two grey alleys of variable length which terminated in two grey goal boxes **(a)**. Choosing white led to a goal box with food; choosing black led to an empty box. Grice found that learning was noticeably impaired by as short a delay as 0.5 s, and severely impaired by 5 s, with little learning at a delay of 10 s **(b, c)**. This deficit could be ameliorated by having more discriminable (black and white) goal boxes, or forcing the rats to make discriminable motor responses (climbing an incline or dodging between blocks) in the black and white start alleys (data not shown). Figures from Grice (1948).

served to be systematically impaired as the outcome is delayed (Skinner, 1938; Perin, 1943; Grice, 1948; Harker, 1956; Lattal & Gleeson, 1990; Dickinson *et al.*, 1992) (Figure 1). Despite this, normal rats have been shown to acquire free-operant responding (Figure 2) with programmed response–reinforcer delays of up to 32 s, or even 64 s if the subjects are pre-exposed to the learning environment (Dickinson *et al.*, 1992) (Figure 3). Delays do reduce the asymptotic level of responding (Dickinson *et al.*, 1992), though the reason for this is not clear. There are several psychological reasons why action–outcome delays might impair learning or performance of an instrumental response (Ainslie, 1975; Cardinal *et al.*, 2004). As discussed above, it may be that when subjects learn a response with a substantial response–reinforcer delay, they never succeed in representing the instrumental action–outcome contingency fully. Alternatively, they may value the delayed reinforcer less. Finally, the delay may also retard the acquisition of a procedural stimulus–response habit and this might account for the decrease in asymptotic responding. It is presently not known whether responses acquired with delayed reinforcement are governed by a different balance of habits and goal-directed actions than responses acquired with immediate reinforcement.

(a) perfect action–outcome contingency, zero delay
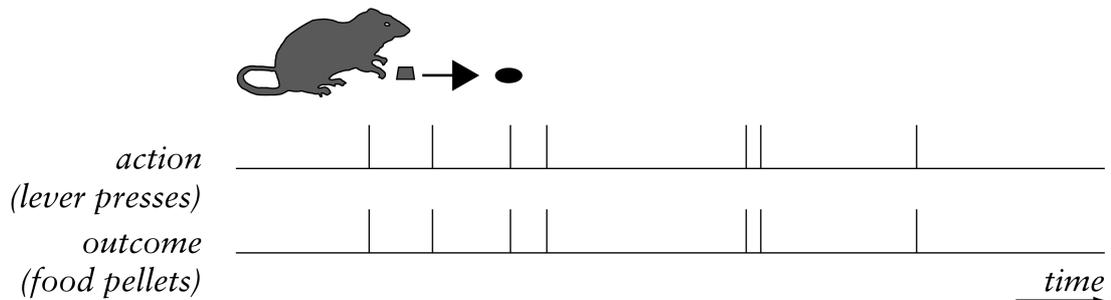
(b) perfect action–outcome contingency, delay > 0
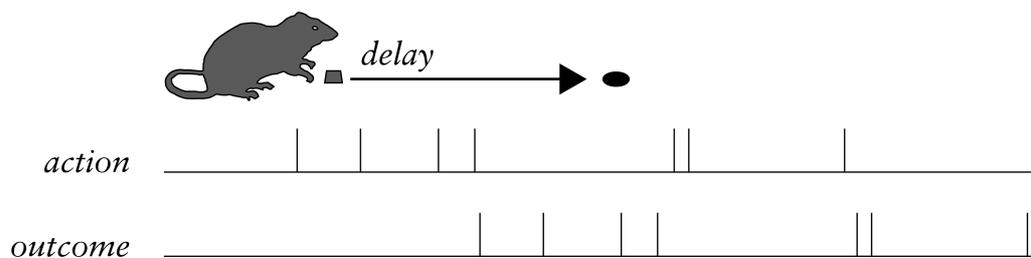
Figure 2: Free-operant learning with delayed reinforcement
When an animal is free to perform an action (operant) to obtain a rewarding outcome, it readily learns to do so if the action–outcome contingency (the increase in the likelihood of obtaining the outcome that is produced by performing the action) is good and if there is no delay between action and outcome **(a)**. Even with a perfect action–outcome contingency, learning is impaired by imposing delays between the action and the outcome **(b)**, yet animals do succeed in this task.
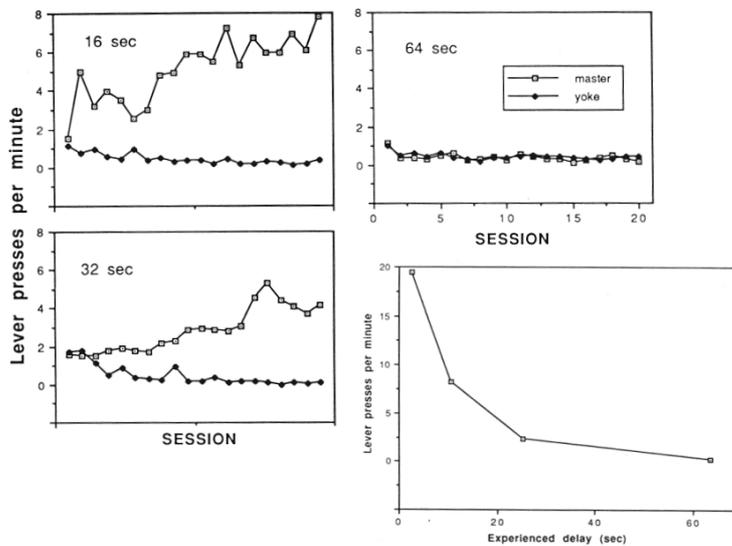
Figure 3: The speed of free-operant learning with delayed reinforcement in normal rats

Dickinson *et al.* (1992) trained rats on a free-operant, fixed-ratio-1 (FR-1) schedule of reinforcement with delays between pressing the lever and obtaining reinforcement (see Figure 2). Responding was compared to that of a yoked control group (who received the same pattern of reinforcement as the "master" rats but whose lever presses had no consequence). The rate of learning, and the asymptotic level of responding, declined across groups as the response–reinforcer delay was increased from 0 to 32 s; rats trained with a 64-s delay failed to learn at all, compared to yoked controls. However, when rats were exposed to the training context, in the absence of the lever or any reinforcers, prior to training, their learning was improved, and successful discrimination was seen even with a delay of 64 s (data not shown), attributed to an underlying process of contextual competition (see text). From Dickinson *et al.* (1992).

1.6.2 Cues and context

Two additional factors must be considered. Cues or signals present during the delay to the reinforcer may become associated with the primary reinforcer, becoming conditioned reinforcers capable of reinforcing actions themselves; conditioned reinforcers may therefore help to bridge action–outcome delays. Indeed, such signals tend to increase responding for delayed reinforcers (Lattal, 1987; Mazur, 1997). One other important factor in learning to act using delayed reinforcement may be the role of the environmental context. The animal's task is to attribute the outcome to its actions; instead, it may erroneously associate the outcome with the context, since the context is a cue that is temporally closer to the outcome than the action is. The longer the delay, the more this contextual competition comes to impair the learning of the action–outcome contingency. Instrumental conditioning with delayed reinforcement can be enhanced if rats are exposed to the relevant contextual cues prior to instrumental training, and this enhancement is lessened if "free" (non-contingent) rewards are given during the contextual pre-exposure periods (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994). These results are consistent with the theory that during the action–outcome delay, contextual cues compete with the action to become associated with the outcome; pre-exposing the animals to the context with no consequences reduces this contextual competition, by making the context a bad predictor of the outcome (perhaps via latent inhibition or learned irrelevance), and this in turn makes the action–outcome contingency more salient and easier to learn (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994).

1.7 CHOICE WITH DELAYED AND UNCERTAIN REINFORCEMENT IN NORMAL ANIMALS

1.7.1 Delayed and probabilistic reinforcement: equivalent or distinct processes?

It has been suggested that delay (or temporal) discounting, the process by which delayed reinforcers lose value, and probability (or odds) discounting, the process by which uncertain reinforcers lose value, reflect the same underlying process (Rotter, 1954; Mischel, 1966; Rachlin *et al.*, 1986; Stevenson, 1986; Rachlin *et al.*, 1987; Mazur, 1989; Rachlin *et al.*, 1991; Mazur, 1995; Green & Myerson, 1996; Mazur, 1997; Sozou, 1998). For example, choosing an uncertain reinforcer five times but only obtaining it on the fifth response might be seen as equivalent to a very long delay, on average, between choice of the reinforcer and its eventual delivery. Alternatively, delays may be seen as entailing the ecological risk of losing the reward during the delay. In animal models, while subjects are learning to respond for delayed or probabilistic rewards, both may initially be similarly unpredictable (although delayed rewards can become more accurately predicted following learning in a manner that stochastic rewards cannot). However, there is evidence that time and probability discounting are different and dissociable processes (Ho *et al.*, 1999; Mitchell, 2003; Green & Myerson, 2004). Most simply, it is not surprising that currency inflation affects human decisions involving delayed but not probabilistic financial reward (Ostaszewski *et al.*, 1998). Moreover, the absolute magnitude of rewards can have different effects on delayed and probabilistic discounting (Green *et al.*, 1999; Myerson *et al.*, 2003; Green & Myerson, 2004). A study looking at human choices in a gambling task found that individuals' propensity to choose rapidly (one, perhaps motoric, measure of delay aversion) and their propensity to bet large amounts of money on uncertain outcomes (a measure of risk taking) represented independent factors (Deakin *et al.*, 2004). Some studies have found abnormal delay discounting, but not uncertainty discounting, in drug addicts (Vuchinich & Calamas, 1997; Mitchell, 1999; Mitchell, 2003; Reynolds *et al.*, 2004b), while gamblers have been observed to discount probabilistic rewards less steeply than controls (i.e. to take risks) without showing differences in delay discounting (Holt *et al.*, 2003).

1.7.2 Temporal or delay discounting

In a typical delayed reinforcement choice task, a subject chooses between an immediate, small ("smaller, sooner" or SS) reward or a large, delayed ("larger, later" or LL) reward; the temporal discounting function quantifies the effect of the delay on preference (Figure 4). Early models of choice assumed an exponential model of temporal discounting (see Kacelnik, 1997a), so that if V_0 is the value of a reinforcer delivered immediately, then the value of a reinforcer delivered after time t is

$$V_t = V_0 e^{-Kt}$$

where K quantifies an individual's tendency to "discount" the future (to value delayed rewards less). The exponential model makes intuitive sense, whether you consider the underlying process to be one in which the subject has a constant probability of "forgetting" its original response per unit time (making it progressively less available for reinforcement), one in which the "strength" of the response's representation decays to a certain proportion of its previous value at each time step, or one in which the subject behaves as if there is a constant probability of losing the delayed reward per unit of waiting time. A S–R learning 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
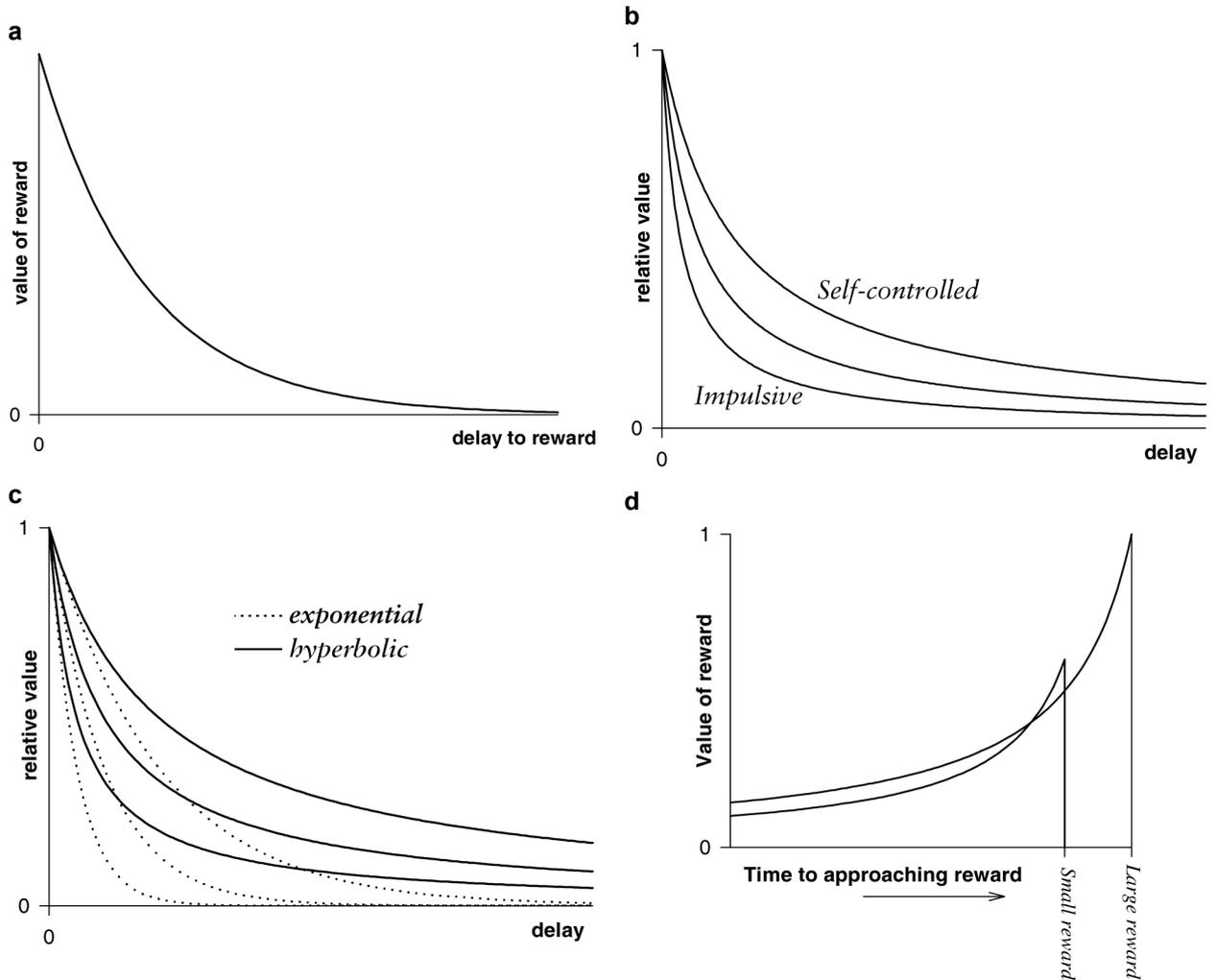
Figure 4: Temporal discounting
(a) The basic, intuitive, and well-validated phenomenon of temporal discounting is that the subjective value of a reward declines monotonically as the reward is progressively delayed: all other things being equal, immediate rewards are worth more than delayed rewards. **(b)** Individuals may vary in their propensity to discount delayed rewards. Individuals who discount the future steeply are said to be impulsive; individuals who discount the future shallowly (giving the future relatively greater weight) are said to be self-controlled. **(c)** Different mathematical models of temporal discounting have been proposed; exponential and hyperbolic discounting are shown. Exponential temporal discounting is described by the equation *value = immediate value* × e$^{-K \cdot delay}$. Hyperbolic temporal discounting is governed by the equation *value = immediate value / (1 + K·delay)*. Large values of the discounting parameter *K* give the steepest curve (the most impulsive behaviour) in both cases. There is strong empirical support for the hyperbolic, not the exponential, discounting model. One critical difference in the predictions of these two models is the phenomenon of preference reversal, since hyperbolic discounting allows curves for different rewards to cross. **(d)** Preference reversal, illustrated for two hypothetical rewards. 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 may come a time just before delivery of the small reward when the value of the small reward exceeds that of the large reward; preference reverses and the small reward is chosen. Figures adapted from Ainslie (1975) (and also published in Robbins *et al.*, 2005).

strength at the previous instant. If x_t is the trace strength at time t and A is the starting value, then

$$x_t = x_0 e^{-kt}$$

and

$$x_{t+1} = e^{-k} x_t$$

However, the exponential model has been emphatically rejected by experimental work with humans and other animals. Instead, temporal discounting appears to follow a hyperbolic or very similar discount function, such as

$$V = \frac{V_0}{1 + Kt}$$

(Grice, 1948; Mazur, 1987; Mazur *et al.*, 1987; Grace, 1996; Richards *et al.*, 1997b).

One interesting prediction that emerges from hyperbolic (but not exponential) models 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 4d), and such preference reversal is a reliable and important experimental finding (see Ainslie, 1975; Green *et al.*, 1981; Bradshaw & Szabadi, 1992; Ainslie, 2001). For example, humans generally prefer £100 now to £200 in three years' time, but also generally prefer £200 in nine years to £100 in six years, despite this being the same choice viewed at a different time. If you are aware that your preference may change in this way, you may be able to improve your happiness in the long run by using self-control strategies (Ainslie & Monterosso, 2003). Ainslie (2001) refers to this as bargaining with your future self; the most famous example of precommitment is that of Odysseus (Homer, ~800 BC / 1996, book 12, translation lines 44–60 and 172–217) (Figure 5); others include the use of disulfiram by alcoholics and social precommitment by announcing publicly one's intention to diet. Even pigeons have been observed to use the self-control strategy of precommitment. When pigeons choose between SS and LL rewards, they are often impulsive (Rachlin & Green, 1972), choosing the SS reward, but they have been shown to work to avoid being offered the option of choosing the SS alternative (Ainslie, 1974; Ainslie & Herrnstein, 1981).

It is also worth noting that in the hyperbolic discounting model and all others in which preference reversal occurs, the value at any one moment cannot be calculated directly from the value immediately preceding it in time; therefore, hyperbolic discounting implies that more information is being maintained by the agent than is required for exponential discounting.

It is not known why hyperbolic discounting arises (Kacelnik, 1997a), or what neuropsychological processes are responsible for it. Such discounting might in principle result from some combination of poor knowledge of the contingencies between actions and their outcomes at long delays, weak S–R habits, or because subjects are perfectly aware that the delayed reward is available but assign a low value to it (Cardinal *et al.*, 2003b). Hyperbolic discounting might also be explicable as the overall effect of two or more different systems—for example, a cognitive, declarative system that exhibits minimal or exponential discounting, plus phenomena such as PIT, conditioned salience, or "visceral factors" that make rewards more salient and promote their choice when they are immediately available (Loewenstein, 1996; Cardinal *et al.*, 2003b; Gjelsvik, 2003; Loewenstein & O'Donoghue, 2004). As discussed above, perhaps the most obvious difference between studies of human impulsive choice and animal models is that humans can be offered explicit choices (hypothetical or real: the difference does not appear to be important; Lagorio & Madden, 2005) without prior experience of the situation (Rachlin *et al.*, 1991; Myerson & Green, 1995; de Wit *et al.*, 2002)—"pre-packaged" action–outcome contingencies. Other animals must learn these contingencies through experience, implying that the whole gamut of psychological representations that contribute to their actions (including goal-directed actions, S–R habits, and conditioned reinforcers) can influence their choices. Nevertheless, hyperbolic discounting has been observed in humans and other experimental animals.

a

b

Figure 5: An early example of precommitment

(a) Waterhouse (1891), *Ulysses and the Sirens*, depicting Odysseus (Ulysses) from Homer's (~800 BC / 1996) *Odyssey* (book 12, translation lines 44–60 and 172–217). Aware that the Sirens—originally bird-like creatures in Greek mythology—would lure his ship onto the rocks through the eldritch influence of their song on men's minds, yet wishing to hear their song for himself, Odysseus commands his men to stop up their ears and lash him to the mast. He gives them strict instructions not to untie him until they are safely past the sirens, and to ignore any further instructions from him until that point. **(b)** Later painters depicted the sirens in more human fashion, or as mermaids; this is Draper's (1909) painting of the same title.

1.7.3 Uncertainty discounting

Similarly, the dominant model of uncertainty or probability discounting (Rachlin *et al.*, 1986; Rachlin *et al.*, 1991; Ho *et al.*, 1999; Green & Myerson, 2004) suggests that subjects calculate a value for each reinforcer, according to its size and other parameters, and discount this by multiplying it by $1/(1+H\theta)$. In this equation, θ represents the odds against obtaining the reinforcer, $\theta = (1 − p)/p$, where p is the probability of obtaining the reward, and H represents an odds discounting parameter that is specific to the individual subject but stable over time for that subject. In this model, value is a hyperbolic function of the odds θ:

$$V = \frac{magnitude}{1 + H \cdot \theta}$$

Such a hyperbolic function is supported by empirical research, at least in humans (Rachlin *et al.*, 1986; Rachlin *et al.*, 1991; Rachlin & Siegel, 1994; Kacelnik, 1997b; Richards *et al.*, 1999b; Rachlin *et al.*, 2000). Preference reversal effects are also observed in choice under risk or uncertainty (Slovic & Lichtenstein, 1983; Lopes, 1994), with subjects preferring gambles with a low probability of winning a large prize when asked to assign monetary values to the gambles, but then preferring gambles with moderate probabilities and prizes when faced with a direct choice—that is, the task used to measure preference alters that preference. Ho *et al.* (1999) suggested that hyperbolic processes of discounting apply to the delay, probability (odds), and magnitude of a reward, and that these three discounting processes are independent, multiplicative, and each governed by its own discounting parameter (*K* for delay, *H* for probability/odds, *Q* for magnitude) that is relatively stable for an individual. Their combined model is therefore as follows:

$$V = \frac{1}{1 + K \cdot delay} \times \frac{1}{1 + H \cdot \theta} \times \frac{magnitude}{magnitude + Q}$$

It should be noted in passing that although effects of delay, probability, magnitude, and so forth are often assumed to be calculated independently (Killeen, 1972; Rachlin *et al.*, 1991; Ho *et al.*, 1999), and though there is some support for this assumption (Mazur, 1987; Mazur, 1997), others have found that the effects of reinforcer delay and magnitude are not independent (Ito, 1985; White & Pipe, 1987). In addition, as discussed below in the context of drug addiction, humans may show quantitatively different temporal (delay) discounting for qualitatively different reinforcers, such as drugs and money. Furthermore, deprivation of one commodity can selectively increase preference for SS over LL rewards for that commodity (e.g. Mitchell, 2004a), suggesting that parameters such as *K* and/or *Q* are not unitary parameters that apply to all reinforcers, and/or that additional parameters specific to reinforcer classes must be added to characterize behaviour fully.

1.8 SYSTEMIC PHARMACOLOGICAL STUDIES OF DELAYED OR UNCERTAIN REINFORCEMENT

Given the importance of impulsive choice in disorders such as addiction (Poulos *et al.*, 1995; Heyman, 1996; Bickel *et al.*, 1999; Evenden, 1999a; Mitchell, 1999) and ADHD (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998), a number of groups have studied the effects on impulsive choice of manipulating neurochemical and neuroanatomical systems implicated in these disorders. I will review pharmacological and neurochemical studies first. To date, more have examined choice involving delayed reinforcement than choice involving uncertain reinforcement, and many more have used appetitive positive reinforcement (reward) than aversive reinforcement such as punishment.

1.8.1 Serotonin (5-HT)

Serotonin (5-hydroxytryptamine, 5-HT) has long been implicated in impulse control. Drugs that suppress 5-HT function were observed to reduce behavioural inhibition, making animals more impulsive in a "motor" sense, as defined above (Soubrié, 1986; Evenden, 1999b). Correlational studies have indicated that low cerebrospinal fluid (CSF) levels of the 5-HT metabolite 5-hydroxyindoleacetic acid (5-HIAA)

are associated with risk taking in monkeys (Mehlman *et al.*, 1994; Evenden, 1998) and impulsive aggression, violence, and suicide in humans (Åsberg *et al.*, 1976; Linnoila *et al.*, 1983; Brown & Linnoila, 1990; Linnoila *et al.*, 1993; Mann, 2003).

Forebrain 5-HT depletion leads to impulsive choice in a variety of paradigms (Wogar *et al.*, 1993; Richards & Seiden, 1995; Bizot *et al.*, 1999; Mobini *et al.*, 2000b) and has been suggested to steepen the temporal discounting function, such that delayed rewards lose their capacity to motivate or reinforce behaviour (Wogar *et al.*, 1993; Ho *et al.*, 1999; Mobini *et al.*, 2000a). The 5-HT-depleted animal becomes hypersensitive to delays, or hyposensitive to delayed reward. As delayed rewards have unusually low value, the animal chooses SS rewards over LL rewards, a characteristic of impulsivity (Ainslie, 1975). Conversely, increasing 5-HT function with the 5-HT indirect agonist fenfluramine decreases impulsive choice (Poulos *et al.*, 1996). Since choice between SS and LL rewards may be affected by changes in the sensitivity to reinforcer magnitude as well as reinforcer delay (Ho *et al.*, 1999), it is important to note that 5-HT depletion does not appear to alter reinforcer magnitude discrimination (Mobini *et al.*, 2000a; Mobini *et al.*, 2000b).

Altered 5-HT function has also been strongly implicated in depression (see e.g. Delgado *et al.*, 1990; Feldman *et al.*, 1997, pp. 842–847; Caspi *et al.*, 2003), but the relationship between depression, impulsivity, and 5-HT is complex. The precise neurochemical abnormality or set of abnormalities in depression is far from clear (e.g. Feldman *et al.*, 1997; Dhaenen, 2001; Stockmeier, 2003). There is no clear-cut relationship between depression itself and levels of 5-HIAA in the CSF (Åsberg, 1997; Feldman *et al.*, 1997, p. 843), although antidepressant drugs themselves tend to lower CSF 5-HIAA (see Bäckman *et al.*, 2000). However, there is a consistent association between low CSF 5-HIAA and suicidal behaviour—not only in depression, but also in schizophrenia and other disorders (see Träskman-Bendz *et al.*, 1986; Cooper *et al.*, 1992; Åsberg, 1997; Cremniter *et al.*, 1999). Patients who are prone to suicide, many of whom are depressed, show high impulsivity (Plutchik & Van Praag, 1989; Apter *et al.*, 1993; Corruble *et al.*, 2003). Thus, low 5-HT function has been linked with impulsive behaviour, which is a risk factor for suicide, and abnormalities of the 5-HT system are also associated with depression, also a strong risk factor for suicide.

However, the results relating 5-HT to impulsivity are not wholly clear-cut. The effects of forebrain 5-HT depletion to promote impulsive choice have sometimes been transient (Bizot *et al.*, 1999) or not observed (Winstanley *et al.*, 2003), and a nonselective 5-HT antagonist has been observed to promote self-controlled choice (Evenden & Ryan, 1996). In humans, lowing 5-HT levels via dietary tryptophan depletion (Biggio *et al.*, 1974; Clemens *et al.*, 1980; Delgado *et al.*, 1989) decreases levels of 5-HT metabolites in CSF (Carpenter *et al.*, 1998; Williams *et al.*, 1999), an indirect indicator of brain 5-HT levels. However, although tryptophan depletion may increase "motor" impulsivity in some tasks (Walderhaug *et al.*, 2002), it does not affect stop-signal reaction time (Clark *et al.*, 2005; Cools *et al.*, 2005), a basic measure of motor control, and it has not been shown to increase impulsive choice in humans (Crean *et al.*, 2002). Likewise, a recent rodent study found that forebrain 5-HT depletion increased motor impulsivity but not delay discounting (Winstanley *et al.*, 2004a). Furthermore, 5-HT efflux in prefrontal cortex (PFC), as measured by microdialysis (as opposed to CSF metabolite levels or whole-tissue post mortem measurement) centred on the prelimbic cortex (PrL), was unexpectedly found to be positively correlated with premature responding in an attentional task, a form of motor impulsivity (Dalley *et al.*, 2002). Post mortem analysis of the same subjects failed to show differences in total tissue 5-HT or 5-HIAA levels between the more impulsive and more self-controlled subgroups. 5-HT may modulate impulsivity in different ways depending on the involvement of different receptor subtypes (Evenden, 1999b; Evenden & Ryan, 1999; Winstanley *et al.*, 2004c). Furthermore, the acute effects of serotonergic drugs on impulsivity

can be the opposite of the chronic effects (Liu *et al.*, 2004), with evidence for complex adaptations within the PFC 5-HT system.

Although manipulations of 5-HT have influenced choice involving delayed reinforcement, there is less evidence that they influence choice involving uncertainty and risk. Although forebrain 5-HT depletion has affected temporal (delay) discounting, as discussed above, it does not appear to influence choice involving probabilistic reinforcement. Dietary tryptophan depletion has not been shown to affect probability discounting in humans (Anderson *et al.*, 2003; Rogers *et al.*, 2003; but see Cools *et al.*, 2005); similarly, forebrain 5-HT depletion in rats does not affect choice between small, certain rewards and large, uncertain rewards (Mobini *et al.*, 2000b).

1.8.2 Noradrenaline (NA)

Relatively little is known about the role of noradrenaline (NA) in delayed or probabilistic reinforcement. It has been suggested that NA neurons encode some aspects of uncertainty in the general sense of making predictions in a given context, in a manner complementary to that of acetylcholine (ACh) (Yu & Dayan, 2005). In causal studies, systemic NA blockade has been shown to affect decision making under uncertainty in humans, by reducing the discrimination between magnitudes of different losses when the probability of losing was high (Rogers *et al.*, 2004a), though NA reuptake inhibition has not been shown to affect the Iowa gambling task (O'Carroll & Papps, 2003), in which subjects must choose between decks of cards differing in magnitude and probability of their expected gains and losses (Bechara *et al.*, 1994).

1.8.3 Dopamine (DA)

1.8.3.1 Temporal difference learning and dopamine

Since prediction of the future is of key importance in designing artificial intelligence agents, a number of mathematical and computational models have been developed to learn from delayed and/or probabilistic reinforcement (Russell & Norvig, 1995), including some forms of Q-learning (Watkins, 1989) and temporal difference (TD) learning (Sutton, 1988). Some models have been compared directly to mammalian neural systems. For example, the TD learning model of Sutton (1988) has been extended to an actor–critic architecture (see Barto, 1995; Houk *et al.*, 1995). In this scheme, a "critic" has access to sensory and motor information and primary reinforcement, and learns to predict reward on the basis of this information using a TD algorithm. "Immediate" reinforcement is held to follow the causing action by one time unit, and the reinforcement at time t is referred to as r_t. Delayed reinforcement is given a lesser weighting by being multiplied by a factor γ for every time step it is delayed (where $0 \leq \gamma < 1$); high γ indicates a strategic or long-term orientation and low γ indicates a tactical, short-term, or impulsive orientation. If the critic is perfect, then its prediction P would be

$$P_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots$$

Therefore, the prediction for time t–1 would be

$$P_{t-1} = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots$$

and thus, for perfect prediction,

$$P_{t-1} = r_t + \gamma(r_{t+1} + \gamma r_{t+2} + \dots)$$

$$P_{t-1} = r_t + \gamma P_t$$

$$r_t + \gamma P_t - P_{t-1} = 0$$

The TD error δ can therefore be defined as

$$\delta = r_t + \gamma P_t - P_{t-1}$$

This quantity δ represents the difference between predicted and actual reward. The critic learns by adjusting its reinforcement prediction on the basis of the TD error: if $\delta > 0$, reward occurred that was not predicted, and the prediction at $t–1$ should be increased for next time; if $\delta < 0$, reward was predicted but did not occur, and the prediction at $t–1$ should be decreased. The critic teaches not only itself but also an "actor", which selects an action, and then modifies the propensity to perform that action on the basis of the TD error (if $\delta = 0$, the consequences of the last action were expected; if $\delta > 0$, the consequences were better than expected, and the response tendency of the action made at $t–1$ should be strengthened; if $\delta < 0$, the consequences were worse than expected, and the response tendency at $t–1$ should be decreased).

The result is that if a consistent sequence of stimuli predicts reward, this system will learn the sequence, with the TD error teaching the system about earlier and earlier consistent predictors with each iteration. As the critic learns about future rewards, it is able to teach the actor to act on the basis of them. Thus the system exemplifies S–R learning with an enhanced ability to act on the basis of future reward. It has been of particular neurobiological interest since the firing of midbrain dopamine (DA) neurons appears to correspond very closely to the TD error δ (Schultz *et al.*, 1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Daw & Touretzky, 2002; McClure *et al.*, 2003b; Schultz, 2006), and other components of the basal ganglia innervated by midbrain DA neurons have been proposed to correspond to the actor and critic, be those components the matrix and striosome compartments of the striatum (Houk *et al.*, 1995) or the dorsal and ventral striatum (O'Doherty *et al.*, 2004).

1.8.3.2 Psychostimulants and impulsivity

However, the original interest in the relationship between DA and impulsivity stems 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 DA from storage vesicles in the terminals of dopaminergic neurons, and prevent DA re-uptake from the synaptic cleft, potentiating its action (for references see Feldman *et al.*, 1997, pp. 293/552/558). It has been proposed that many features of ADHD, including preference for immediate reinforcement and hyperactivity on simple reinforcement schedules, are due to abnormally steep temporal discounting, and that this is due to a hypofunctional nucleus accumbens (Acb) DA system (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998; Johansen *et al.*, 2002). Indeed, they go on to suggest Acb DA as the specific culprit (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998). Acb DA has long been implicated in aspects of responding for reinforcement, though its role is not yet fully understood (Cardinal *et al.*, 2002a; Salamone *et al.*, 2005). However, whether ADHD is characterized by a hypodopaminergic or a hyperdopaminergic state, and how this and other (e.g. noradrenergic/serotonergic) abnormalities might be "normalized" by psychostimulants is controversial (Swanson *et al.*, 1998; Zhuang *et al.*, 2001; Seeman & Madras, 2002; Solanto, 2002; Fone & Nutt, 2005; Russell *et al.*, 2005; Williams & Dayan, 2005).

Many of the inferences regarding the neural abnormalities in children with ADHD have 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; Russell *et al.*, 2005). This rat exhibits pervasive hyperactivity and attention problems that resemble ADHD, exhibits a steeper "scallop" of responding on fixed-interval (FI) schedules of reinforcement, which can be interpreted as abnormally high sensitivity to immediate reinforcement (Sagvolden *et al.*, 1992), is impulsive on measures of "execution impulsivity" (Evenden & Meyerson, 1999), and has a complex pattern of abnormalities in its DA system (de Villiers *et al.*, 1995; Russell *et al.*, 1995; Papa *et al.*, 1996; Carey *et al.*, 1998; Papa *et al.*, 1998; Russell *et al.*, 1998; Russell, 2000). Depolarization- and psychostimulant-induced DA release in Acb 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 DA system projecting to the Acb (de Villiers *et al.*, 1995; Russell *et al.*, 1998; Russell, 2000), though abnormalities have also been found in DA release in slices of dorsal striatum and PFC (Russell *et al.*, 1995). Within the Acb, differences in gene expression and DA 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 may reflect a lack of effectiveness of delayed reinforcement, and has been suggested to underlie at least some subtypes of ADHD (Sagvolden *et al.*, 1998; Sagvolden & Sergeant, 1998; Kuntsi *et al.*, 2001; Sonuga-Barke, 2002). The efficacy of psychomotor stimulants in ADHD (Bradley, 1937; Solanto, 1998) suggests that they might promote the choice of delayed rewards. In fact, the effects of acute administration of psychostimulants on laboratory models of impulsive choice have varied. Some studies have found that they promote choice of delayed reinforcers (Sagvolden *et al.*, 1992; Richards *et al.*, 1997a; Richards *et al.*, 1999a; Wade *et al.*, 2000; de Wit *et al.*, 2002), while others have found the opposite effect (Logue *et al.*, 1992; Charrier & Thiébot, 1996; Evenden & Ryan, 1996); the same psychostimulant can even have opposite effects in different tasks designed to measure impulsivity (Richards *et al.*, 1997a). One factor that may explain some of these discrepant effects is the presence of cues or signals present during the delay to the larger/later alternative. Such signals tend to increase responding for the delayed reinforcer (Lattal, 1987; Mazur, 1997), perhaps because they become associated with the primary reinforcer and themselves become conditioned reinforcers, thus affecting choice (Williams & Dunn, 1991). Psychostimulants increase the effect of conditioned reinforcers (Hill, 1970; Robbins, 1976; Robbins, 1978; Robbins *et al.*, 1983), and their effects in delayed reinforcement choice tasks can depend on whether explicit signals are presented during the delay (Cardinal *et al.*, 2000). However, conditioned reinforcement is certainly not the only procedural difference between studies that have found differing effects of psychostimulants.

1.8.3.3 Dopamine D$_1$ and D$_2$ receptors and impulsivity

It should also be emphasized that few studies of the effects of psychostimulants on impulsive choice have addressed the pharmacological basis of their effects. Some of the effects may indeed not be dopaminergic: for example, the effects of amphetamine may depend in part on 5-HT (Winstanley *et al.*, 2003). However, Wade *et al.* (2000) have shown that D$_2$-type DA receptor antagonists and mixed D$_1$/D$_2$ antagonists induce impulsive choice, while D$_1$-type receptor antagonists do not, suggesting that DA D$_2$ receptors normally promote choice of delayed reinforcement.

The role of DA in reward uncertainty is also not well understood. DA neurons respond to reward prediction errors with changes in their phasic firing rate, as discussed above, and may also carry information in their sustained firing rate specifically about reward uncertainty (Fiorillo *et al.*, 2003; Fiorillo *et al.*, 2005; Niv *et al.*, 2005; Tobler *et al.*, 2005), but little is known of the causal role of DA in choice involving uncertain rewards.

1.8.4 Relationship between addictive drugs and impulsivity

Given that impulsivity is part of the syndrome of drug addiction, with impulsive choice playing a prominent role in maintaining the selection of drugs of abuse in favour of other, longer-term rewards (Poulos *et al.*, 1995; Heyman, 1996; Bickel *et al.*, 1999; Evenden, 1999a; Mitchell, 1999; APA, 2000), the relationship between addictive drugs and impulsive choice is of clear interest. Studies examining discounting in addicts have focused primarily on delay, rather than uncertainty discounting (see Mitchell, 2003; 2004b; 2004a). There is little evidence for differences in uncertainty discounting among smokers (Mitchell, 1999) or alcohol abusers (Vuchinich & Calamas, 1997), though alcohol has been shown to modify decision making under uncertainty (George *et al.*, 2005). However, decision-making deficits in risk-taking and gambling tasks have been demonstrated in opiate and amphetamine users (Rogers *et al.*, 1999a; Ersche *et al.*, 2005; Leland & Paulus, 2005). The fact that the deficits were in some cases correlated with the number of years of abuse suggests (but does not prove) that the deficits were drug-induced; the possibility remains that the decision-making deficits predated and predisposed to the addiction.

Abnormally steep delay discounting has been demonstrated in drug addicts, including alcoholics (Vuchinich & Calamas, 1997; Petry, 2001), cocaine users (Coffey *et al.*, 2003; Kirby & Petry, 2004), opiate users (Madden *et al.*, 1997; Kirby *et al.*, 1999; Kirby & Petry, 2004), and smokers (Bickel *et al.*, 1999; Mitchell, 1999; Mitchell, 2003; Reynolds *et al.*, 2004b; Ohmura *et al.*, 2005; Reynolds, 2006); again, the question of cause and effect is hard to determine, although steeper discounting in current addicts compared to ex-addicts again raises the possibility of an effect of ongoing drug use. Many studies have looked at the pharmacological effects of addictive drugs on measures of impulsivity including response inhibition; rather fewer have looked specifically at delay and/or probability discounting in a formal experimental (causal) design. Chronic cocaine administration transiently increases delay discounting (increases impulsive choice) in rats (Paine *et al.*, 2003), as does acute morphine administration (Kieres *et al.*, 2004); acute administration of psychostimulants was discussed above, and chronic methamphetamine has been shown to increase impulsive choice in rats (Richards *et al.*, 1999a). In keeping with everyday experience, alcohol has been observed to increase risk taking (Lane *et al.*, 2004). However, the findings for a given drug have not always been consistent. For example, Ortner *et al.* (2003) recently found that alcohol reduced delay discounting in humans, while Richards *et al*. (1999b) found no effect of alcohol on this measure; in contrast, several investigators have found impulsive choice to be induced by alcohol in rats (Tomie *et al.*, 1998; Evenden & Ryan, 1999; Hellemans *et al.*, 2005) and some also in humans (Reynolds *et al.*, 2006). Benzodiazepines have not been shown to affect impulsive choice in humans (Reynolds *et al.*, 2004a), and in different studies have been observed both to increase (Thiebot *et al.*, 1985; Cardinal *et al.*, 2000) and to decrease (Evenden & Ryan, 1996) impulsive choice in rats. These discrepancies may in some cases be due to the sensitivity of the particular task used, but may also be because the drugs (or the state of addiction) do not have a unitary effect on discounting, but one which depends heavily on the situation and the particular choices involved. For example, Mitchell has shown that cigarette deprivation increases choice impulsivity when decisions concern cigarettes, but not when they concern money (Mitchell, 2004a); likewise, smokers temporally discount cigarettes more than money (Bickel *et al.*, 1999), as well as discounting money more than controls; opiate abusers discount opiates more than money (Madden *et al.*, 1999) (Figure 6), and cocaine users discount cocaine more than money (Coffey *et al.*, 2003).
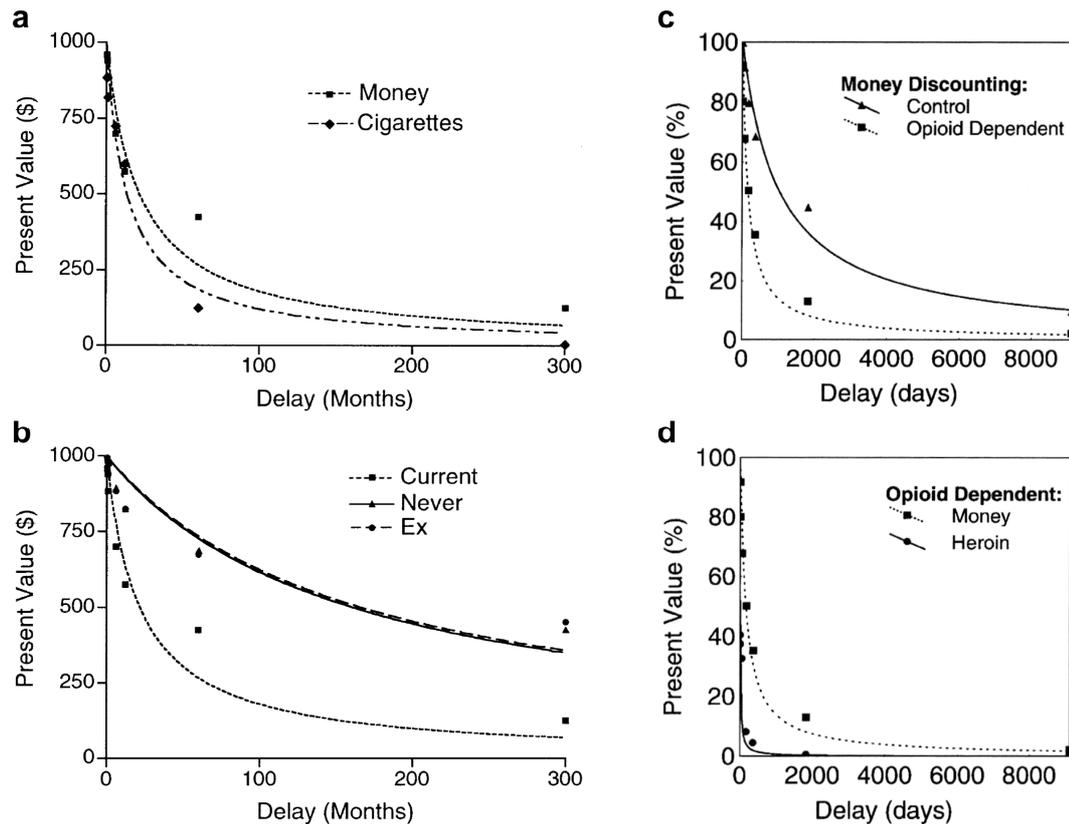
Figure 6: Exaggerated temporal discounting in drug addicts
Summary of a number of studies conducted by Bickel and colleagues examining temporal discounting in human drug addicts. **(a)** Smokers discount cigarettes slightly more steeply than money. **(b)** Current smokers discount money more than ex-smokers, or those who have never smoked, do. **(c)** Opioid-dependent humans discount money more than control subjects do. **(d)** Opioid addicts discount heroin more steeply than they discount money. Data for smokers from Bickel *et al.* (1999); data for heroin addicts from Madden *et al.* (1999), reproduced from Bickel & Johnson (2003).

1.9 ANATOMY AND CONNECTIONS OF KEY LIMBIC STRUCTURES

In this section I will outline the concept of the limbic system and in particular the anatomy of the limbic corticostriatal "loop", the striatal component of which is the nucleus accumbens. I will describe the basic anatomy and connectivity of the Acb and of the hippocampus, the two areas whose roles in delayed/uncertain reinforcement are examined in this thesis.

1.9.1 Anatomy of the limbic corticostriatal "loop"

Early investigations of the functions of hypothalamic regions (Hetherington & Ranson, 1939; Anand & Brobeck, 1951) demonstrated that electrolytic regions of the lateral hypothalamus appeared to leave animals demotivated, with impairments in unlearned behaviour (including aphagia, adipsia, and a reduction in sexual, exploratory, and maternal behaviours) and in learned behaviour (impaired instrumental responding). However, such lesions also disrupt the medial forebrain bundle, a fibre tract that passes through the lateral hypothalamus and contains the projection from midbrain DA nuclei, the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA), to the forebrain. Use of the DA-depleting toxin 6-hydroxydopamine (6-OHDA) showed that lesions of this projection or DA depletion of the striatum, one of its targets, produced a similar pattern of behavioural impairment (Stricker & Zig-

mond, 1976; Marshall & Teitelbaum, 1977); this focused attention on the role of DA and the structures that it innervated in motivated behaviour.

The basal ganglia comprise a number of subcortical nuclei, including the striatum. The striatum may be considered the "input layer" of the basal ganglia; nearly the entire neocortex projects to it (Kemp & Powell, 1971). In turn, the striatum projects to the globus pallidus, which projects via thalamic nuclei back to the cortex; the whole makes up a "loop". It is a particular characteristic of basal ganglia–thalamo-cortical ("corticostriatal") loops that although large areas of cortex send information into the loop, only a relatively small area of cortex is targeted by the return projection. Information flow in different loops is segregated—that is, the loops operate in parallel—and the loops are named for the areas of cortex to which they project. They are the *motor* loop (projecting in primates to the premotor cortex, supplementary motor area, and primary motor cortex, and involved in the initiation of motor acts); the *oculomotor* loop (projecting to the frontal eye fields); the *dorsolateral prefrontal* or "cognitive" loop; the *lateral orbito-frontal* loop, and the anterior cingulate or *limbic* loop (projecting to the anterior cingulate cortex and medial orbitofrontal cortex) (DeLong & Georgopoulos, 1981; Alexander *et al.*, 1986). Indeed, functional segregation (parallel processing) is apparent even within each loop (Alexander & Crutcher, 1990). The loops may also be differentiated on the basis of the parts of the basal ganglia and thalamus they pass through; thus, while inputs to the motor and cognitive loops target the dorsal striatum (caudate–putamen or neostriatum), information entering the limbic loop does so through the ventral striatum. The ventral striatum consists of the Acb, ventromedial portions of the caudate and putamen, and the olfactory tubercle; the largest component is the Acb. Within each corticostriatal loop, the basic circuitry is similar across the dorsal striatum and much of the ventral striatum (Heimer *et al.*, 1995); it is therefore likely that the various basal ganglia loops process information in qualitatively similar ways, with the nature of the cortical target determining the apparent function of each loop.

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

Here, I will focus on the limbic loop, depicted in Figure 7. Its components include many of the structures considered part of the limbic system. The term "limbic" was coined by Broca (1878) for the cortical structures encircling the upper brain stem (limbus, L. edge or border). The "limbic lobe" was suggested to have a role in emotional experience and expression by Papez (1937), concepts later to be elaborated by MacLean (1949; 1952; 1993), who introduced the expression "limbic system" to refer to the limbic lobe and its connections with the brainstem. The limbic system is not precisely defined: as the limbic lobe was considered the neural substrate for emotions, structures whose functions have to do with motivation and emotion have since been added to the anatomical definition. A modern definition of the limbic system in

primates would include cingulate and orbitofrontal cortex; the hippocampal formation, parahippocampal gyrus and mammillary bodies; anterior and medial thalamic nuclei; the nucleus accumbens and ventral pallidum; the amygdala and the hypothalamus. Key elements of the limbic corticostriatal loop are shown in Figure 7 (p. 21), with anatomically accurate depictions in Figure 8 (coronal views; p. 22), Figure 9 (sagittal views; p. 23), Figure 10 (horizontal views; p. 24), and Figure 11 ("glass brain" views; p. 25).
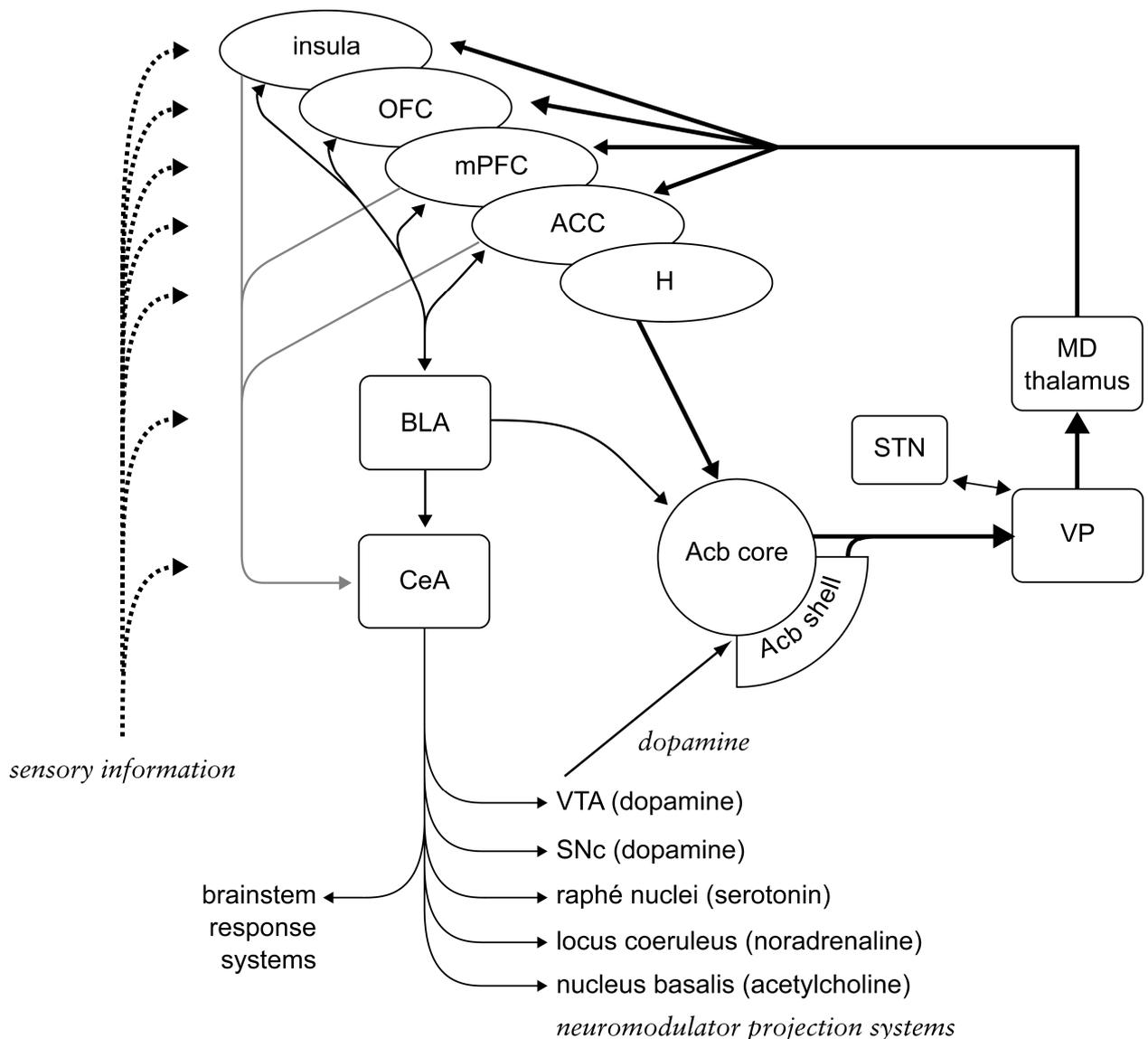
Figure 7: Key elements of the limbic corticostriatal "loop"

Simplified schematic of the limbic corticostriatal loop (after Cardinal *et al.*, 2002a), showing key structures. OFC, orbitofrontal cortex; mPFC, medial prefrontal cortex (prelimbic/infralimbic cortex in the rat); ACC, anterior cingulate cortex; H, hippocampal formation; BLA, basolateral amygdala; CeA, central nucleus of the amygdala; Acb, nucleus accumbens; STN, subthalamic nucleus; VP, ventral pallidum; MD, mediodorsal; VTA, ventral tegmental area; SNc, substantia nigra pars compacta. Not all structures and connections are shown; for example, there are projections from prefrontal cortical regions, including the OFC, to the STN (Berendse & Groenewegen, 1991; Maurice *et al.*, 1998; Hamani *et al.*, 2004).
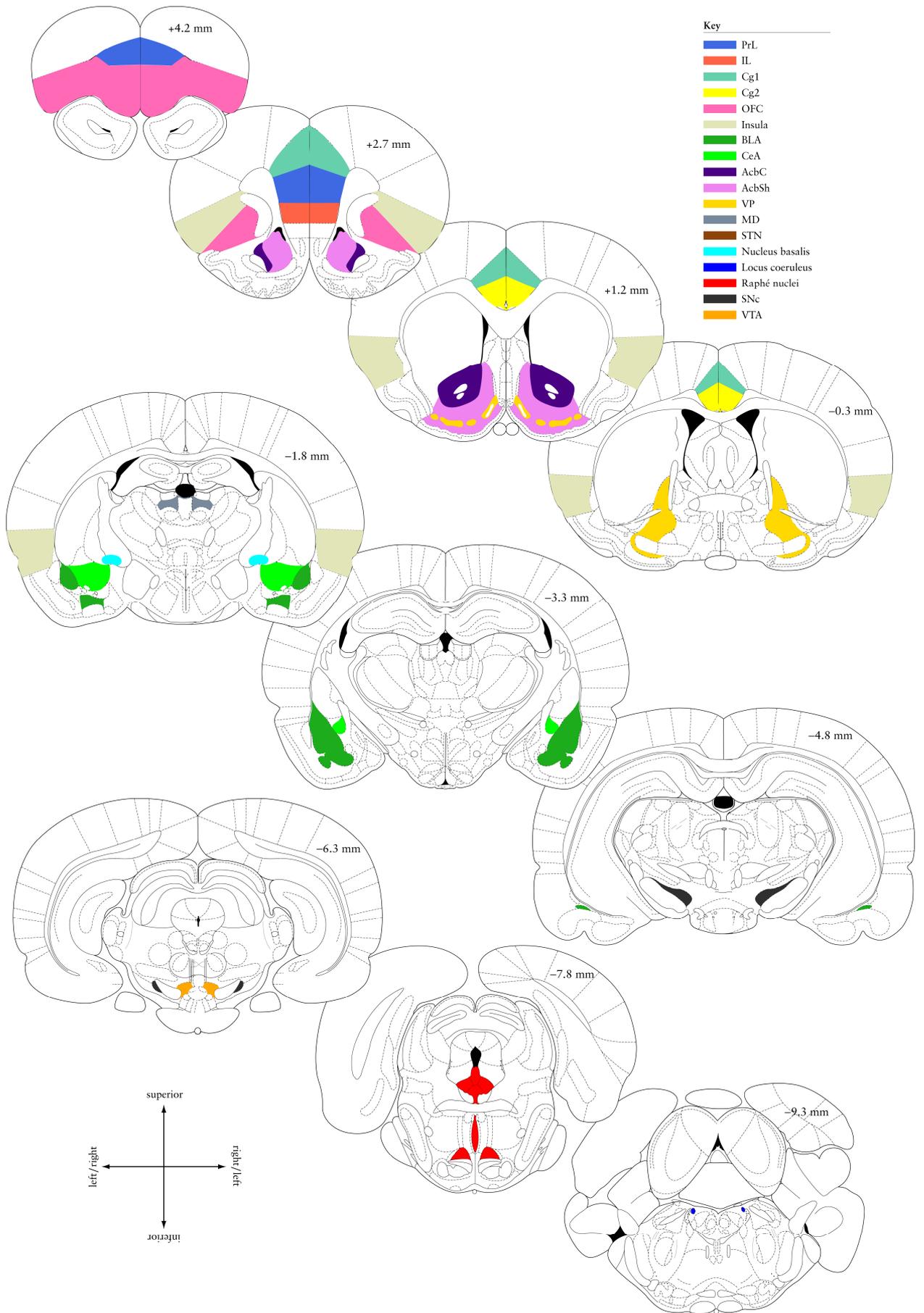
Figure 8: Coronal sections of the rat brain, showing selected limbic and related structures.
For full legend, see p. 26.
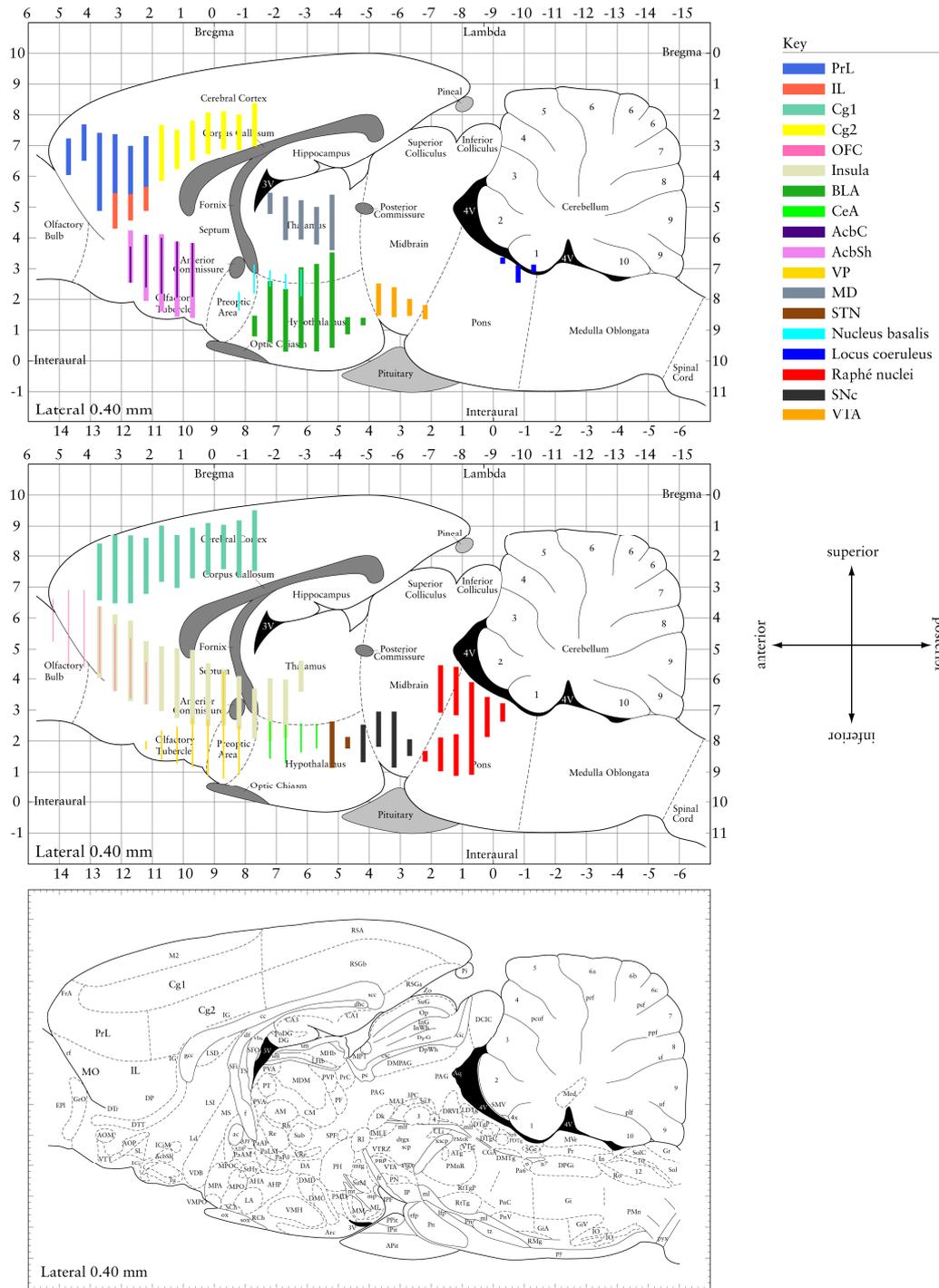
Figure 9: Sagittal paramedian views of the rat brain, showing selected limbic and related structures.
For full legend, see p. 26.
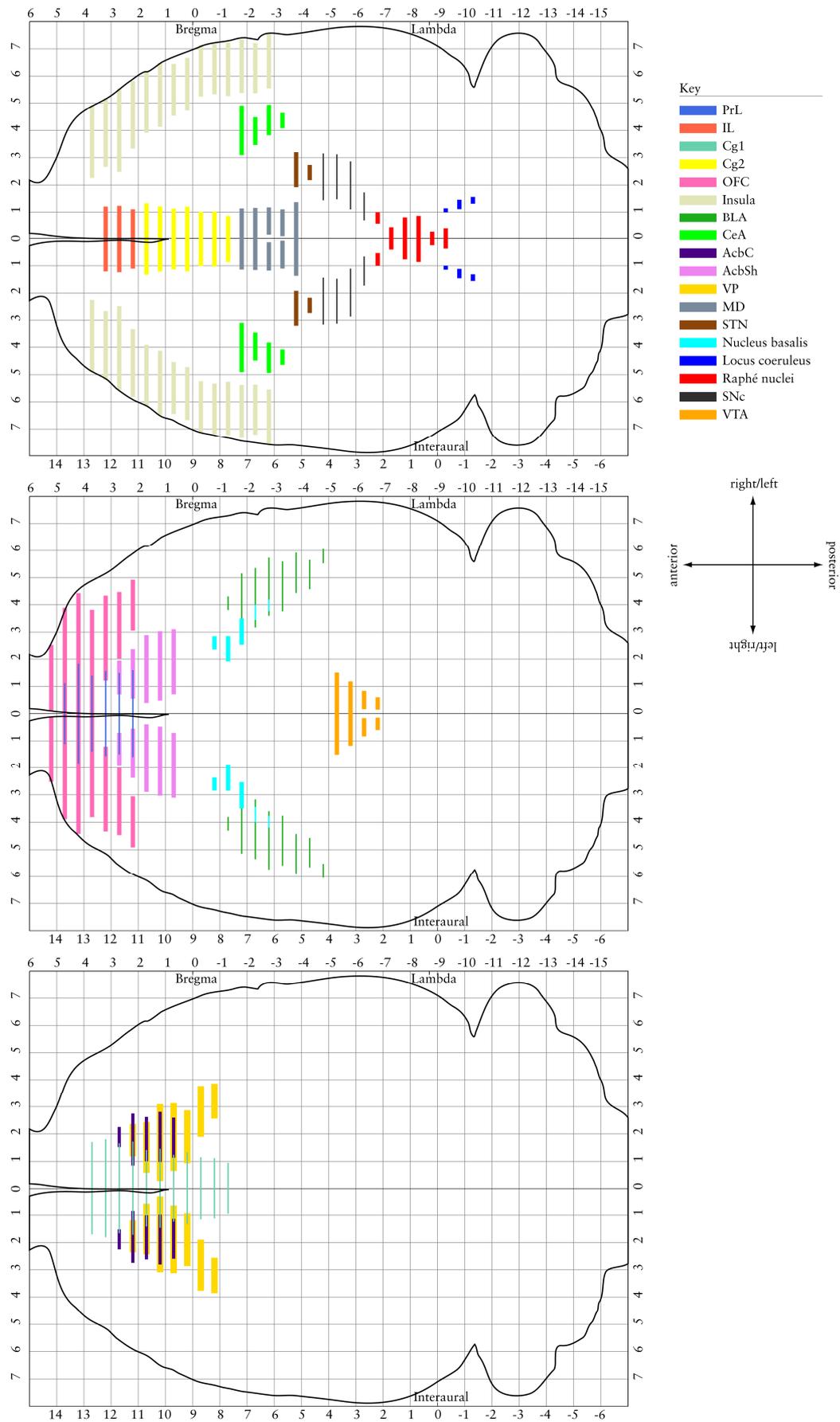
Figure 10: Horizontal views of the rat brain, showing selected limbic and related structures.
For full legend, see p. 26.
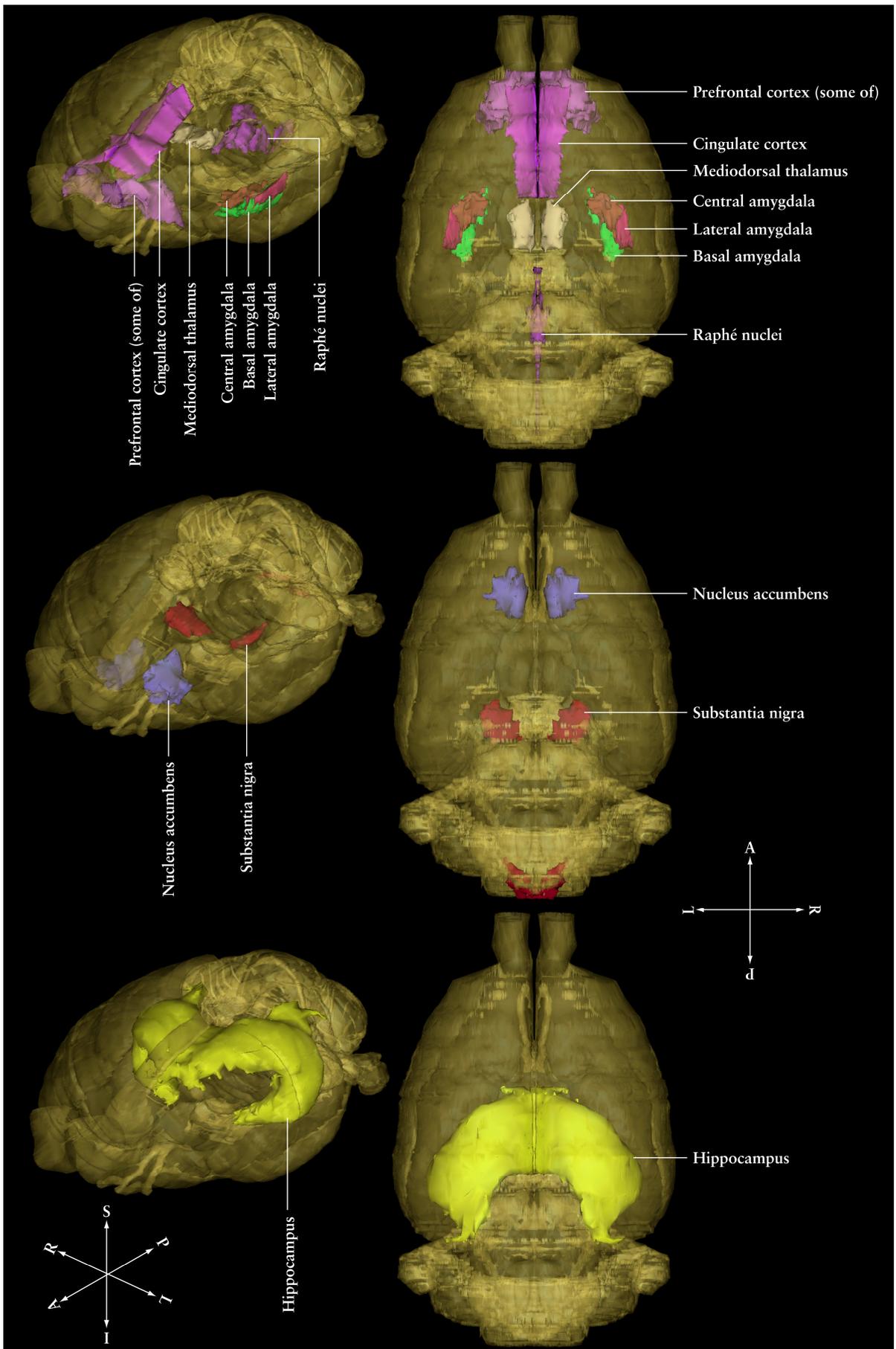
Figure 11: "Glass brain" views showing selected limbic and related structures.
For full legend, see p. 26.

Legends continued from Figure 8–Figure 11 (pp. 22–25).

Figure 8 (p. 22): Coronal sections of the rat brain, showing selected limbic and related structures.

Coronal sections are shown at 1.5 mm intervals (from +4.2 mm to –9.3 mm relative to bregma, with positive being anterior). Colours indicate selected regions of interest. Ventricles are shown in black; the SNc is shown in near-black and is located inferiorly and bilaterally at –4.8 mm and –6.3 mm slices. Coronal sections are taken from Paxinos & Watson (1997). The hippocampus is not highlighted (for which, see Figure 12). **Abbreviations and regional definitions:** PrL, prelimbic cortex; IL, infralimbic cortex; Cg1, cingulate area 1; Cg2, cingulate area 2; OFC, orbitofrontal cortex (including areas MO [medial orbital cortex], VO [ventral orbital cortex], and LO [lateral orbital cortex]). The insula, or insular cortex, includes agranular insular cortex (AI), dysgranular insular cortex (DI), and granular insular cortex (GI). BLA, basolateral amygdalar complex (including the basolateral amygdaloid nucleus BL, the basomedial amygdaloid nucleus BM, and the lateral amygdaloid nucleus La); CeA, central amygdaloid nucleus; AcbC, nucleus accumbens core; AcbSh, nucleus accumbens shell; VP, ventral pallidum; MD, mediodorsal nucleus of the thalamus; STN, subthalamic nucleus (STh in the terminology of Paxinos & Watson, 1997). "Nucleus basalis" refers to the nucleus basalis magnocellularis, or basal nucleus of Meynert (B in the terminology of Paxinos & Watson, 1997). Raphé nuclei shown include the dorsal raphé nucleus (DR), the median raphé nucleus (MnR), and the B9 group of serotonergic cells. SNc, substantia nigra pars compacta (SNC in the terminology of Paxinos & Watson, 1997); VTA, ventral tegmental area.

Figure 9 (p. 23): Sagittal paramedian views of the rat brain, showing selected limbic and related structures.

Top and middle panels: sagittal sections at 0.4 mm lateral to the midline, from Paxinos & Watson (1997), with scales in mm. Superimposed upon these sections are coloured strips at 0.5 mm intervals indicating the maximum vertical extent of each structure at that anteroposterior level, throughout the brain (not just at 0.4 mm lateral to the midline); data from coronal sections of Paxinos & Watson (1997). The groupings of structures in the upper and middle panels are not functional groupings, but are chosen to minimize overlap in the figures. **Bottom panel:** another sagittal view at 0.4 mm lateral to the midline, from Paxinos & Watson (1998), labelled. Not all abbreviations will be defined here; the picture illustrates the surface topography of prefrontal cortex (particularly areas PrL, IL, MO, Cg1, and Cg2) near the midline. Abbreviations are as in Figure 8. The hippocampus is not shown (for which, see Figure 12).

Figure 10 (p. 24): Horizontal views of the rat brain, showing selected limbic and related structures.

Horizontal sections of the rat brain, with scales in mm. The outline of the rat brain is traced from that of *MIVA* version 0.9 (Sullivan & Zhang, 2005). Superimposed upon these are coloured strips at 0.5 mm intervals indicating the maximum horizontal extent of each structure at that anteroposterior level. The groupings of structures in the three panels are not functional, but are chosen to minimize overlap in the figures. The hippocampus is not shown (for which, see Figure 12).

Figure 11 (p. 25): "Glass brain" views showing selected limbic and related structures.

Images showing views of a transparent rat brain shell ("glass brain") in two three-dimensional projections (seen from the front/left/above, and seen from directly above) with structures illustrated as defined in and rendered by *MIVA* version 0.9 (Sullivan & Zhang, 2005). S, superior; I, inferior; A, anterior; P, posterior; R, right; L, left.

1.9.2 Basic anatomy of the nucleus accumbens

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

nections of the Acb are summarized in Table 1 and Table 2. As a recipient of information from a considerable array of limbic structures that projects additionally to nuclei known to be involved in behavioural expression, the Acb has famously been suggested to represent a "limbic–motor interface" (Mogenson *et al.*, 1980).

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

Table 1: Some inputs to the nucleus accumbens

Subcortical connections are nearly all reciprocal. Information from Brog *et al.* (1993) and Berendse *et al.* (1992). See Brog *et al.* (1993) for further details of thalamic connections. Table reproduced from Cardinal (2001).

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

Table 2: Some outputs from the nucleus accumbens

For references, see Pennartz *et al.* (1994). Table reproduced from Cardinal (2001).

1.9.3 Basic anatomy of the hippocampus

The term "hippocampus" is usually taken to mean the cornu ammonis (CA, or Ammon's horn), the dentate gyrus, and the subiculum (Aggleton & Brown, 1999). The cornu ammonis has a number of subfields, termed CA1–4. The hippocampus is archicortex. It has bidirectional links with adjacent entorhinal cortex, which itself communicates with perirhinal and parahippocampal cortex. The other main conduit of information to and from the hippocampus is via the fornix, a fibre tract that starts with its fimbriae (L. fringes)

on the hippocampus, and terminates predominantly in the mammillary bodies (part of the hypothalamus), and the anterior thalamic nuclei, but also in the nucleus accumbens. The mammillary bodies themselves project to these thalamic nuclei via the mammillothalamic tract. The macroscopic anatomy of the rat hippocampus is shown in Figure 12.

Within the hippocampus, there is a well-described trisynaptic circuit (Andersen *et al.*, 1969; 1971) (Figure 13, Figure 14). All major association areas of cortex project reciprocally to the entorhinal cortex. (1) Entorhinal cortex cells project via the *perforant path* directly to the dentate gyrus, crossing the hippocampal fissure in the process. (2) Dentate gyrus cells (specifically, granule cells) project via so-called *mossy fibres* to CA3. (3) In addition to sending axons out along the fornix, CA3 cells project via *Schaffer collaterals* to the CA1 field. After this, CA1 axons project either back to the subiculum (and from there back to entorhinal cortex) or to the fornix. The complete set of circuitry is, of course, more complex than
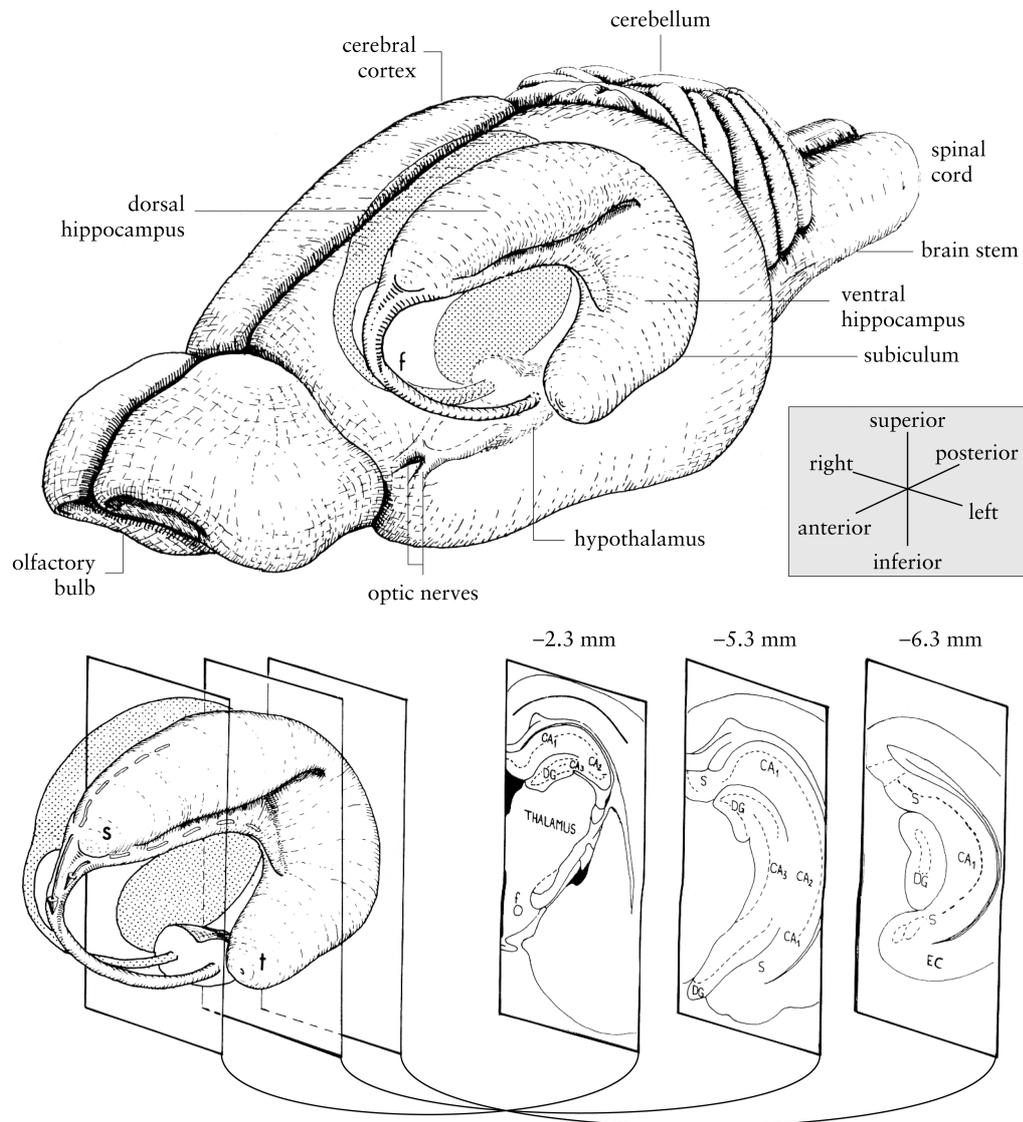
Figure 12: Diagram of the rat hippocampus

Drawings of the rat brain showing the three-dimensional organization of the hippocampus and related structures. Three coronal sections through the left hippocampus are shown at the bottom right of the figure, with their approximate anteroposterior coordinate relative to bregma. CA1, CA2, CA3: cornu ammonis fields 1–3; DG: dentate gyrus; EC: entorhinal cortex; f: fornix; s: septal pole of the hippocampus; S: subiculum; t: temporal pole of the hippocampus. Adapted from Figure 1 of Amaral & Witter (1995); copyright Elsevier 1995; reproduced in Cheung & Cardinal (2005) with permission from Elsevier.

this basic description (Figure 13). The hippocampus also receives important modulatory input, including ACh. The main forebrain cholinergic innervation comes from the nucleus basalis, which provides ACh to neocortex, and the nearby septum (septal nuclei) and diagonal band of Broca in the basal forebrain, which together provide much of the ACh input to the hippocampus. Cholinergic cells of the medial septum project via the fornix to all regions of the hippocampus; in turn, CA3 projects back to the lateral septum, where inhibitory interneurons project to the medial septum.
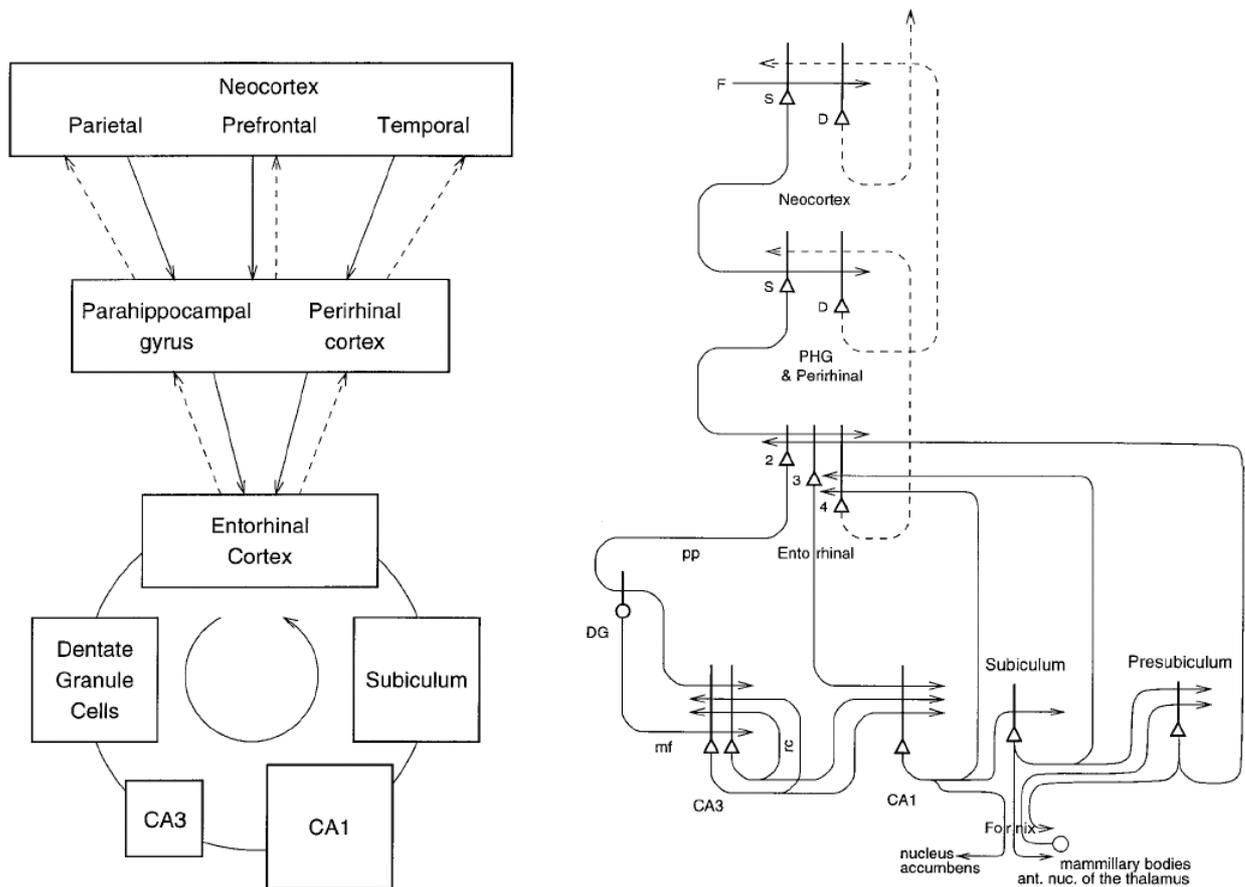
Figure 13: Outline of connections of the hippocampus

The hippocampus comprises the cornu ammonis, the dentate gyrus, and the subiculum. **Left:** basic intrinsic and extrinsic connections of the hippocampus, excluding the fornix. **Right:** synaptic connections in more detail, showing principal excitatory neurons and including fornical connections. CA, cornu ammonis; D, deep pyramidal cells; DG, dentate gyrus granule cells; F, forward inputs to association cortex areas from preceding cortical areas in the hierarchy shown at left; mf, mossy fibres; PHG, parahippocampal gyrus and perirhinal cortex; pp, perforant path; rc, recurrent collateral of the CA3 hippocampal pyramidal cells; S, superficial pyramidal cells; 2, pyramidal cells in layer 2 of the entorhinal cortex; 3, pyramidal cells in layer 3 of the entorhinal cortex. Thick lines above cell bodies represent dendrites. Figure taken from Rolls (2000).
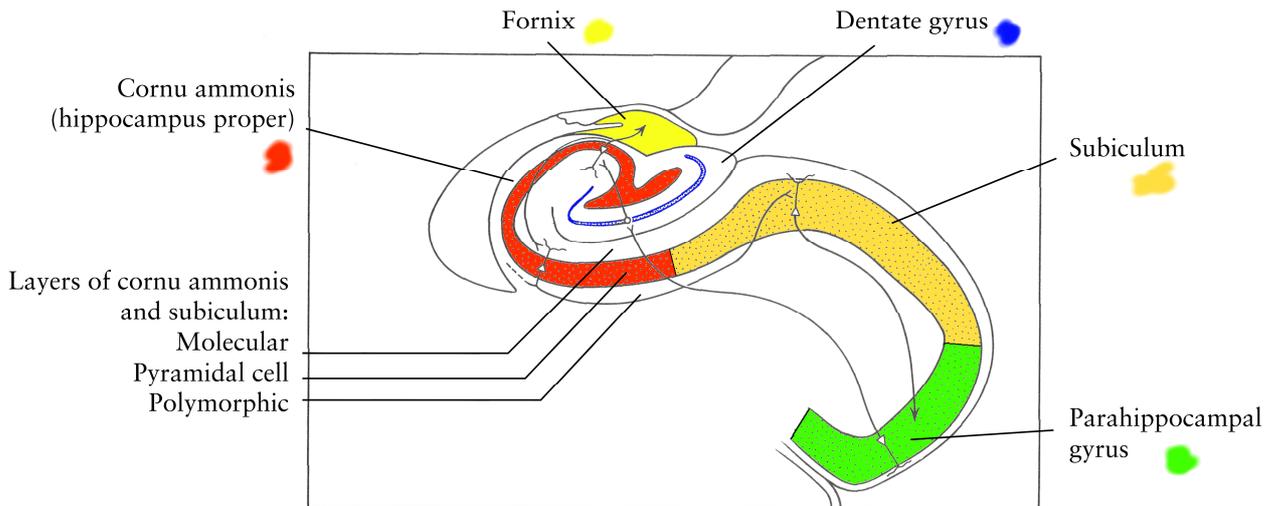
Figure 14: Cross-sectional structure of the primate hippocampus

Coronal section: medial is to the right, superior is upwards. The diagram shows the interlocking C shapes (CA fields and dentate gyrus) that typify the hippocampus, and the major pathways through the hippocampal formation. Modified from Martin (1989, p. 391). CA subfields are numbered with CA1 closest to the subiculum. In this diagram, the entorhinal cortex is considered part of the parahippocampal gyrus.

1.9.4 A note on the interpretation of excitotoxic lesion methods

Although correlative techniques such as electrophysiology and functional neuroimaging allow the functioning of the normal brain to be measured, interventional techniques (such as lesion studies or drug infusions) are required to establish a causal link between a neural structure and an aspect of behaviour. In such studies, the anatomical specificity of the method is important. The use of aspirative or radiofrequency lesions, or local anaesthetic inactivation, will destroy or inactivate neurons in the target area, but will also affect fibres (axons) passing through the target structure, potentially affecting the function of neurons whose cell bodies are elsewhere. In the present thesis, excitotoxic lesion techniques are used to affect neurons in the target site selectively. Excitotoxins typically activate *N*-methyl-D-aspartate-type (NMDA-type) glutamate receptors on neurons, leading to abnormal Ca^{2+} influx and cell death via apoptosis or excitotoxic necrosis; reviews have been provided by Choi (1988; 1995). Table 3 shows the conclu-

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

Table 3: Interpretation of lesion studies

Modified from Cardinal (2001).

sions that may be drawn from some of these interventional techniques.

1.10 BASIC NEUROBIOLOGY OF REINFORCEMENT LEARNING

Having outlined the psychological processes known to play a part in reinforcement learning (p. 2) and the structures of the limbic corticostriatal loop (p. 19), it will be helpful to attempt a rough correspondence before examining the specific contribution of limbic structures in the context of delayed and uncertain reinforcement. A number of limbic cortical and subcortical structures play a role in assessing the value of reinforcers and of stimuli that predict them, and in actions directed at obtaining those reinforcers or stimuli (Cardinal *et al.*, 2002a); here, I will summarize some major themes in reinforcement neuroscience.

1.10.1 Mesolimbic dopamine and the nucleus accumbens

The discovery that rats would work very hard to stimulate regions of their brain electrically—ICSS (Olds & Milner, 1954)—was historically important. Many sites that support ICSS lie on the path of dopaminergic (DAergic) neurons from the SNc and VTA to limbic sites including the ventral striatum (nucleus accumbens), and ICSS is substantially reduced after Acb DA depletion (Fibiger *et al.*, 1987). Furthermore, the rate at which rats learn an ICSS response is correlated with the degree of potentiation of synapses made by cortical afferents onto striatal neurons, a potentiation that requires DA receptors (Reynolds *et al.*, 2001). The natural idea that follows is that DA "stamps in" stimulus–response connections. Indeed, DA has acute effects to modulate corticostriatal transmission, but it also has lasting effects; most likely, the combination of cortical (presynaptic) and striatal (postsynaptic) activity normally induces long-term depression of corticostriatal synapses, but if the same pattern of activity is paired with a pulse of DA, then the active synapses are strengthened (Reynolds & Wickens, 2002). Natural reinforcers, drugs of abuse, and CSs that predict either, trigger increases in DA release in the Acb (Berridge & Robinson, 1998; Datla *et al.*, 2002; Ito *et al.*, 2002; Carelli & Wightman, 2004; Young, 2004). DA neurons fire to unexpected rewards, or unexpected stimuli that predict reward; that is, they signal reward prediction error (Schultz *et al.*, 1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Schultz, 2006). DA neuron firing may be a teaching signal used for learning about actions that lead to reward (Schultz *et al.*, 1997). The Acb similarly responds to anticipated rewards (Schultz *et al.*, 1992; Miyazaki *et al.*, 1998; Martin & Ono, 2000; Schultz *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; de la Fuente-Fernandez *et al.*, 2002; Cromwell & Schultz, 2003; Elliott *et al.*, 2003; McClure *et al.*, 2003a; Bjork *et al.*, 2004; Zink *et al.*, 2004). Other parameters of DA neuronal firing may signal reward uncertainty (Fiorillo *et al.*, 2003; Schultz, 2004; 2006).

An early suggestion was that Acb DA mediated the pleasurable aspects of reward (Wise, 1981; 1982; 1985), but there is good evidence against this idea. Certainly, DA is released in response to appetitive reinforcers (e.g. Fiorino *et al.*, 1993; Wilson *et al.*, 1995; Schultz *et al.*, 1997; Berridge & Robinson, 1998; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Datla *et al.*, 2002; Ito *et al.*, 2002; Carelli & Wightman, 2004; Young, 2004), intra-Acb DA agonists are reinforcing (Phillips *et al.*, 1994), animals may titrate their drug taking to maintain high Acb DA levels (Pettit & Justice, 1989), and some aspects of naturally reinforced and drug-reinforced responding depend on Acb DA (e.g. Pettit *et al.*, 1984; Caine & Koob, 1994; Baker *et al.*, 1998; Ikemoto & Panksepp, 1999; Dickinson *et al.*, 2000; Parkinson *et al.*, 2002; Salamone & Correa, 2002; Salamone *et al.*, 2003). However, Acb DA does not mediate "pleasure" (Fibiger & Phillips, 1988; Robbins & Everitt, 1992; Berridge & Robinson, 1998; Volkow *et al.*, 1999)—though its release may correlate with activity in other systems that do—and reinforcement operates in its absence (Ettenberg *et al.*, 1982; Pettit *et al.*, 1984). Measured by microdialysis techniques, DA is also

released in response to aversive stimuli, CSs that predict them, and other salient stimuli (see e.g. Salamone, 1994; Horvitz, 2000; Young, 2004), which would be consistent with a more general motivational role. These results are not easy to reconcile with electrophysiological studies of DA neuronal firing, which have generally suggested that firing occurs in response to appetitive but not aversive stimuli. It is possible that DA neurons fire to strong, not mild, aversive events, that the DA response to aversive events is gradual rather than phasic, or that local modulation of DA release in target regions dissociates DA release from the firing of DA neurons (reviewed by Salamone, 1994; Horvitz, 2000; Joseph *et al.*, 2003).

Targets of DA neurons certainly influence instrumental learning and responding. It is not clear what structures learn from the DA teaching signal; these probably include the dorsal striatum and PFC, but much attention has focused on the Acb. Blockade of NMDA glutamate receptors in the AcbC has been shown to retard instrumental learning for food under a variable-ratio-2 (VR-2) schedule (Kelley *et al.*, 1997), as has inhibition or over-stimulation of cyclic-adenosine-monophosphate-dependent protein kinase (protein kinase A; PKA) within the Acb (Baldwin *et al.*, 2002a). Concurrent blockade of NMDA and DA D_1 receptors in the AcbC synergistically prevents learning of a VR-2 schedule (Smith-Roe & Kelley, 2000). Once the response has been learned, subsequent performance on this schedule is not impaired by NMDA receptor blockade within the AcbC (Kelley *et al.*, 1997). Furthermore, infusion of a PKA inhibitor (Baldwin *et al.*, 2002a) or a protein synthesis inhibitor (Hernandez *et al.*, 2002) into the AcbC *after* instrumental training sessions impairs subsequent performance, implying that PKA activity and protein synthesis in the AcbC contribute to the consolidation of instrumental behaviour. Thus, manipulation of the Acb can affect instrumental learning.

However, it is also clear that the Acb is not *required* for simple instrumental conditioning—but it is strongly implicated in providing "extra motivation" for behaviour, especially when such motivation is triggered by Pavlovian CSs, or when reinforcers are delayed or require substantial effort to obtain. Rats with Acb or AcbC lesions acquire lever-press responses on sequences of random ratio (RR) schedules at normal or near-normal levels (Corbit *et al.*, 2001; de Borchgrave *et al.*, 2002) and are fully sensitive to changes in the action–outcome contingency (Balleine & Killcross, 1994; Corbit *et al.*, 2001; de Borchgrave *et al.*, 2002). Thus, the Acb is not critical for goal-directed action (see Cardinal *et al.*, 2002a); rather, it appears to be critical for some aspects of motivation that promote responding for rewards in real-life situations. For example, the Acb plays a role in promoting responding for delayed rewards (Cardinal *et al.*, 2001), to be discussed later, and is required for Pavlovian CSs to provide a motivational boost to responding (Hall *et al.*, 2001; de Borchgrave *et al.*, 2002), i.e. for PIT. PIT has sometimes been termed "wanting" (Wyvell & Berridge, 2000; Wyvell & Berridge, 2001), although the term "wanting" could equally refer to the instrumental incentive value underpinning true goal-directed action. PIT can be further enhanced by injection of amphetamine into the Acb (Wyvell & Berridge, 2000) or by corticotropin-releasing hormone (CRH) acting in the AcbSh (Pecina *et al.*, 2006), and depends on DA (Dickinson *et al.*, 2000), possibly under the control of the central nucleus of the amygdala (CeA) (Hall *et al.*, 2001). CeA control of midbrain dopaminergic systems has also been demonstrated in other tasks (Lee *et al.*, 2005a; Lee *et al.*, 2005b; Holland & Gallagher, 2006). Other motivational effects of Pavlovian CSs also depend on the Acb. CSs serve as goals for behaviours (conditioned reinforcers); although lesions of Acb subregions do not prevent animals responding for conditioned reinforcement entirely (Parkinson *et al.*, 1999a), enhancement of DA neurotransmission within the Acb can boost the efficacy of conditioned reinforcement (Taylor & Robbins, 1984; 1986; Cador *et al.*, 1991; Parkinson *et al.*, 1999a). CSs that have been paired with reward also elicit approach (Brown & Jenkins, 1968); this effect also depends on the Acb (Parkinson *et al.*, 1999a; Parkinson *et al.*, 1999b; Parkinson *et al.*, 2000c) and its DA innervation

(Parkinson *et al.*, 2002). Acb DA may also be involved in learning this approach response, again perhaps under the control of the CeA (Parkinson *et al.*, 2000b; Hall *et al.*, 2001; Cardinal *et al.*, 2002b; Parkinson *et al.*, 2002; Phillips *et al.*, 2003). Acb DA also contributes directly to subjects' motivation to work hard (Ikemoto & Panksepp, 1999; Salamone & Correa, 2002; Salamone *et al.*, 2003). In naturalistic situations, rewards are frequently available only after a delay, require considerable effort to achieve, and are signalled by environmental stimuli; thus, the Acb is central to a number of processes that require motivation (Mogenson *et al.*, 1980).

This motivational process has been suggested to be particularly significant in some addictions, and modification of it may have therapeutic potential. Although DA systems are affected by drugs of abuse and natural reinforcers such as food, some abused drugs may be more potent in this regard. For example, both food and drugs of abuse increase Acb DA, but the DA response to drugs of abuse may not habituate to the same extent as that to food (Di Chiara, 1998; Di Chiara, 2002). Sensitization occurs following psychostimulant administration directly into the VTA, which induces hypersensitivity to DA in the Acb (Cador *et al.*, 1995) and enhances the response to Pavlovian CSs associated with reward (Harmer & Phillips, 1999; Taylor & Horger, 1999; Wyvell & Berridge, 2001). It is not yet clear to what extent sensitization contributes to human addiction (Sax & Strakowski, 2001), but it has been suggested that a sensitized response to drug-associated cues contributes to drug craving—that this "incentive motivational" system becomes sensitized (Robinson & Berridge, 1993). In present animal models, however, drug sensitization enhances responding for food, or responding to CSs for food (Taylor & Horger, 1999; Wyvell & Berridge, 2001; Olausson *et al.*, 2003), but in human addiction, responding for non-drug reinforcement declines relative to that for drug reinforcement (APA, 2000). Amphetamine sensitization also enhances the subsequent development of habits (Nelson & Killcross, 2006). In any case, in animal models of drug-seeking behaviour controlled by drug-associated stimuli (Everitt & Robbins, 2000), lesions of the AcbC or disruption of its glutamatergic neurotransmission reduce drug seeking (Di Ciano & Everitt, 2001; Hutcheson *et al.*, 2001b), probably by reducing the motivational impact of the CSs. DA D3 receptors are particularly concentrated in the Acb and amygdala (Sokoloff *et al.*, 1990), and D3 receptor antagonists (Vorel *et al.*, 2002; Di Ciano *et al.*, 2003) and partial agonists (Pilla *et al.*, 1999; Cervo *et al.*, 2003) reduce cue-controlled cocaine seeking or relapse to cocaine taking in animal models. Some manipulations that reduce drug seeking or reinstatement of drug taking in animal models, such as DA D3 receptor antagonists, do not reduce food seeking in a similar manner (Vorel *et al.*, 2002; Di Ciano *et al.*, 2003).

1.10.2 Habits and the dorsal striatum

The development of motor habits may depend on dorsal striatal plasticity (Packard & McGaugh, 1996), which may in turn depend on DA receptors (Reynolds *et al.*, 2001; Reynolds & Wickens, 2002). Expression of S–R habits requires the dorsal striatum (Packard & McGaugh, 1996; Yin *et al.*, 2004), and the balance between habits and goal-directed behaviour may also be regulated by the prelimbic and infralimbic cortex (Killcross & Coutureau, 2003), subdivisions of the rat PFC. Dorsal striatal DA release is also a correlate of well-established cocaine seeking (Ito *et al.*, 2002).

1.10.3 Action–outcome contingency knowledge, planning and value: the PFC and amygdala

The PFC (specifically, prelimbic cortex) is required for rats to represent the contingencies between actions and their outcomes (Balleine & Dickinson, 1998; Corbit & Balleine, 2003), and acquisition of instrumental responses on a simple schedule is also disrupted synergistically by concurrent blockade of

NMDA and DA D$_1$ receptors in the PFC (Baldwin *et al.*, 2002b). The PFC is also involved in extinction (Myers & Davis, 2002)—the cessation of responding when a CS or response is no longer paired with re-inforcement. Extinction is not "unlearning" but involves the learning of new, inhibitory ("CS → not-US" or "CS → don't respond") associations (see Mackintosh, 1974; Delamater, 2004). Lesions of the ventral medial PFC interfere with the extinction of Pavlovian conditioned freezing in the rat (Morgan *et al.*, 1993; Morgan & LeDoux, 1995; Morgan & LeDoux, 1999). The PFC interacts with the amygdala, an important site of CS–US association in this task (see Davis, 2000; LeDoux, 2000), and may suppress conditioned freezing when it is no longer appropriate (see Garcia *et al.*, 1999; Myers & Davis, 2002; Quirk *et al.*, 2003; Rosenkranz *et al.*, 2003).

The orbitofrontal cortex (OFC) is part of the PFC with a particular role in the assessment of reinforcer value; it has bidirectional connections to the amygdala, and both are heavily implicated in the retrieval of the value of primary reinforcers based on information from CSs (see Cardinal *et al.*, 2002a; Balleine *et al.*, 2003; Lindgren *et al.*, 2003; Pickens *et al.*, 2003; O'Doherty, 2004). The OFC may encode the economic value of goods directly (e.g. Padoa-Schioppa & Assad, 2006). In humans, the OFC and amygdala are also activated during extinction of Pavlovian conditioning (Gottfried & Dolan, 2004). The amygdala regulates the DA signal to the Acb (Everitt *et al.*, 2000; Parkinson *et al.*, 2000a; Hall *et al.*, 2001; Cardinal *et al.*, 2002a; Phillips *et al.*, 2003). Goal-directed action requires that action–outcome contingencies interact with the incentive value of goals (Dickinson, 1994; Dickinson & Balleine, 1994); the connection between the amygdala and the PFC (Pitkänen, 2000) may provide this functional link (Coutureau *et al.*, 2000; Arana *et al.*, 2003; Gottfried *et al.*, 2003; Holland & Gallagher, 2004). The retrieval of incentive value about food reinforcers also requires the gustatory cortex, the insula (Balleine & Dickinson, 1998; Balleine & Dickinson, 2000).

1.10.4 Hedonic assessment

Hedonic assessment of rewards themselves ("liking" or "pleasure"), does not depend on dopaminergic processes (Pecina *et al.*, 1997; Berridge & Robinson, 1998; Dickinson *et al.*, 2000; Pecina *et al.*, 2003). Instead, it involves opioid mechanisms in the AcbSh and other systems in the pallidum and brainstem (Berridge, 2000; Kelley & Berridge, 2002). Intra-Acb μ opioid agonists also affect food preference, increasing the intake of highly palatable foodstuffs including fat, sweet foods, salt, and ethanol (Zhang *et al.*, 1998; Zhang & Kelley, 2000; Kelley *et al.*, 2002; Zhang & Kelley, 2002; Will *et al.*, 2003; Ward *et al.*, 2006). The effect can be two-way, with chronic ingestion of chocolate inducing adaptations in endogenous Acb opioid systems (Kelley *et al.*, 2003).

1.10.5 The hippocampus and the representation of context

Interest in the hippocampus as a memory store stemmed from early observations of human amnesia following medial temporal lobe resection (Scoville & Milner, 1957; Corkin *et al.*, 1997) and many subsequent animal models (for recent reviews, see Squire, 1992; Murray & Mishkin, 1998; Baxter & Murray, 2001a; Baxter & Murray, 2001b; Clark *et al.*, 2001b), together with the discovery of synaptic long-term potentiation (LTP), a cellular mechanism of memory, in the hippocampus (Bliss & Lømo, 1973; Morris, 1994). Human anterograde amnesia has also resulted from damage to diencephalic structures, and it has been suggested that a circuit involving the hippocampus, mammillary bodies, and anterior thalamic nuclei is essential for episodic memory formation (Delay & Brion, 1969; Aggleton & Brown, 1999). There is substantial contemporary debate on the exact type of memory supported by the hippocampus and adjacent cortical regions, and the manner in which they do so (Gaffan & Harrison, 1989; Gaffan, 1992; Morris &

Frey, 1997; Eichenbaum *et al.*, 1999; Griffiths *et al.*, 1999; Morris, 2001; Good, 2002; Day *et al.*, 2003; Fortin *et al.*, 2004). However, there is good evidence that the hippocampus contributes to the representation of context.

The idea that the hippocampus plays a role in contextual representations developed from the original discovery of cells in the rat hippocampus that increased their firing rate when the rat was at a particular location in its environment—"place cells" (O'Keefe & Dostrovsky, 1971). O'Keefe & Nadel (1978) suggested that the hippocampus functions as a "cognitive map", informing the rat where it is in the world (recently reviewed by Eichenbaum *et al.*, 1999). Lesion studies support the idea that the hippocampus is critical in navigation. For example, Morris *et al.* (1982) showed that rats with hippocampal lesions were impaired at a task in which they had to learn the location of a hidden submerged platform in a tank full of opaque liquid, now known as the Morris water maze. The deficit appears to depend on navigating relative to a constellation of cues in the room, as hippocampal lesions do not impair the ability to head in a particular direction to a stimulus that bears a fixed relation to the platform (Pearce *et al.*, 1998). Water maze performance is damaged by dorsal, not ventral hippocampal lesions (Moser *et al.*, 1995). Learning in the water maze can be blocked by the glutamate NMDA receptor antagonist D-(–)-2-amino-5-phosphonopentanoic acid (AP-5), which blocks LTP (Morris *et al.*, 1986); similar effects follow NMDA receptor subunit mutations. However, the effects of AP-5 are attenuated if the rats are trained in a different water maze beforehand (Bannerman *et al.*, 1995), so the role of the NMDA receptors may not be a specifically spatial one. Human imaging studies also support the idea of a role in the hippocampus in navigation (e.g. Maguire *et al.*, 1997; Maguire *et al.*, 1998; Maguire *et al.*, 2000).

It has been argued that the hippocampus doesn't encode a map in the conventional sense; rather, it appears that place cells encode the relationship between subsets of cues in the environment, independent of other cues (Eichenbaum *et al.*, 1999). Hippocampal neurons also encode nonspatial features (Wood *et al.*, 1999). Eichenbaum *et al.* (1999) suggest that the hippocampus can encode spatial information because this is a special case of encoding the relations between stimuli. These relations are useful for navigation when they are spatial relations, but the memories encoded by the hippocampus can be used for other purposes. The use of a more abstract relationship is demonstrated by transitive inference. If a subject learns that B > C (where ">" denotes "should be chosen over") and C > D, then the logical property of transitivity should allow it to infer that B > D. Dusek & Eichenbaum (1997) have shown that fornix transection and perirhinal/entorhinal cortex lesions, both of which partially disconnect the hippocampus, impair transitive inference in rats. Similarly, Eichenbaum and colleagues have demonstrated that the hippocampus contributes to the memory for sequences of events in rats (Fortin *et al.*, 2002; Ergorul & Eichenbaum, 2006).

Both the hypothesis that the hippocampus encodes spatial relationships (e.g. Eichenbaum *et al.*, 1999), and the hypothesis that it encodes visual scenes (Gaffan, 1992), predict that the hippocampus might be involved in associating together the many visual and non-visual elements that make up a spatial environment, or context. Therefore, it might be expected that the hippocampus contributes to contextual conditioning. In a prototypical task, if a rat receives tone–shock pairings in a distinct environment, it may subsequently show "fearful" reactions to the tone (discrete CS conditioning) and also the environment (contextual conditioning). Lesions of the hippocampus have been shown to impair Pavlovian conditioning to a contextual CS, but not to a discrete CS, in rats (Hirsh, 1974; Selden *et al.*, 1991; Kim & Fanselow, 1992; Phillips & LeDoux, 1992; Honey & Good, 1993; Jarrard, 1993; Kim *et al.*, 1993; Phillips & LeDoux, 1994; Phillips & LeDoux, 1995; Chen *et al.*, 1996; Maren & Fanselow, 1997; Anagnostaras *et al.*, 1999; Rudy *et al.*, 2002), at least for some processes involving contextual representation (Good & Honey, 1991;

Holland & Bouton, 1999; Good, 2002). Context-specific neuronal firing patterns develop in the hippocampus of rats required to discriminate different contexts (Smith & Mizumori, 2006). In some cases, discrete CS conditioning has even been enhanced following hippocampal lesions (e.g. Ito *et al.*, 2005), which may reflect a reduction in contextual competition (see p. 8).

1.11 NEUROANATOMICALLY SPECIFIC STUDIES OF DELAYED OR UNCERTAIN REINFORCEMENT

In recent years, a number of studies have examined the effects of focal excitotoxic or neurochemical lesions on choice and learning involving delayed or uncertain rewards, in additional to correlational studies using functional imaging, microdialysis, and electrophysiology. These studies centre on interconnected structures forming part of the limbic corticostriatal loop (Figure 7). Initial work focused on the Acb and two of its cortical afferents, the anterior cingulate cortex (ACC) and the medial prefrontal cortex (mPFC), as structures potentially involved in regulating choice between alternative reinforcers, for three main reasons.

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 & Everitt, 1996; Salamone *et al.*, 1997; Everitt *et al.*, 1999; Parkinson *et al.*, 2000a; Cardinal *et al.*, 2002a). 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.*, 2000c; Cardinal *et al.*, 2003a) and the mPFC (e.g. Balleine & Dickinson, 1998; Richardson & Gratton, 1998; Bechara *et al.*, 1999; Tzschentke, 2000).

Second, these regions are important recipients of dopaminergic and serotonergic afferents (Fallon & Loughlin, 1995; Halliday *et al.*, 1995), and pharmacological manipulations of DA and 5-HT 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 SHR, differences in DA 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 DA release have been detected in the Acb (de Villiers *et al.*, 1995; Russell *et al.*, 1998; Russell, 2000) and PFC (Russell *et al.*, 1995), in addition to possible dysfunction in the dorsal striatum and amygdala (Russell *et al.*, 1995; Papa *et al.*, 2000).

These early studies, described below, indicated a role for the AcbC in choosing delayed rewards; subsequent work has attempted to delineate the contribution of structures connected to it, and these will be reviewed in turn.

1.11.1 Nucleus accumbens core (AcbC)

1.11.1.1 Choice involving delayed reinforcement

The Acb responds to anticipated rewards in a variety of species (Schultz *et al.*, 1992; Miyazaki *et al.*, 1998; Martin & Ono, 2000; Schultz *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; Cromwell & Schultz, 2003; Bjork *et al.*, 2004; Izawa *et al.*, 2005). As discussed above, it is innervated by DA neurons that respond to errors in reward prediction in a manner appropriate for a teaching signal (Schultz *et al.*,

1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Schultz, 2006), interventional studies have shown it to be a key site for the motivational impact of impending rewards (reviewed by Robbins & Everitt, 1996; Salamone *et al.*, 1997; Everitt *et al.*, 1999; Parkinson *et al.*, 2000a; Cardinal *et al.*, 2002a; Robbins *et al.*, 2005), and Acb abnormalities have been observed in rat models of ADHD (de Villiers *et al.*, 1995; Papa *et al.*, 1996; Carey *et al.*, 1998; Papa *et al.*, 1998; Russell *et al.*, 1998; Russell, 2000; Sadile, 2000).

Causal experimental studies have shown that excitotoxic lesions of the AcbC produce impulsive choice, reducing rats' preference for large/delayed rewards, compared to small/immediate rewards (Cardinal *et al.*, 2001; 2003b). These studies used a task in which rats were offered regular choices between a one-pellet immediate reward and a four-pellet reward delayed from 0–60 s (Figure 15). No cues were present during the delay, to avoid any potential confounds arising from conditioned reinforcement effects (Cardinal *et al.*, 2000), and subjects were trained preoperatively, assigned to matched groups, operated upon, and retested postoperatively, to avoid any possible effects of the lesion on learning of the task. AcbC-lesioned subjects (Figure 16) were rendered impulsive in their choices: they exhibited a profound deficit in their ability to choose a delayed reward, and persisted in choosing impulsively even though they were made to experience the larger, delayed alternative at regular intervals (Figure 17). This effect 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.

Although a few lesioned subjects avoided the large-reinforcer alternative postoperatively even when the delay was zero, this was probably due to within-session generalization from trial blocks at which delays were present (Evenden & Ryan, 1996; Cardinal *et al.*, 2000), as prolonged training in the absence of delays restored a near-absolute preference for the large reinforcer in the majority of subjects—who were then much more impulsive than shams again when delays were re-introduced (Figure 18) (Cardinal *et al.*, 2003b). These results indicate that AcbC-lesioned rats were able to discriminate the two reinforcers, but preferred immediate small rewards to larger delayed rewards.

Similar effects on preference are observed following lesions of the caudal lobus parolfactorius in the chick, thought to be the avian counterpart of the Acb (Izawa *et al.*, 2003).

Recently, AcbC lesions have also been found to impair performance on a task requiring rats to choose between an uncertain immediate reward and a certain delayed reward (Pothuizen *et al.*, 2005). One alternative required completion of a fixed-ratio-5 (FR-5) response for immediate delivery of a food pellet with probability $P = 0.25$; the other required completion of an FR-5 response for delayed certain delivery of an identical food pellet. AcbC lesions reduced rats' preference for the delayed, certain alternative, following sustained testing (Pothuizen *et al.*, 2005). AcbC lesions also reduced efficiency (the number of responses made per reward earned) in a differential-reinforcement-of-low-rates (DRL) schedule (Pothuizen *et al.*, 2005), in which animals must respond below a certain rate in order to obtain reward. This is much like the effects of whole-Acb lesions (Reading & Dunnett, 1995), although the DRL task may also be susceptible to changes in general levels of motor activity: AcbC-lesioned rats are hyperactive (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001), and hyperactivity would itself tend to reduce DRL efficiency.
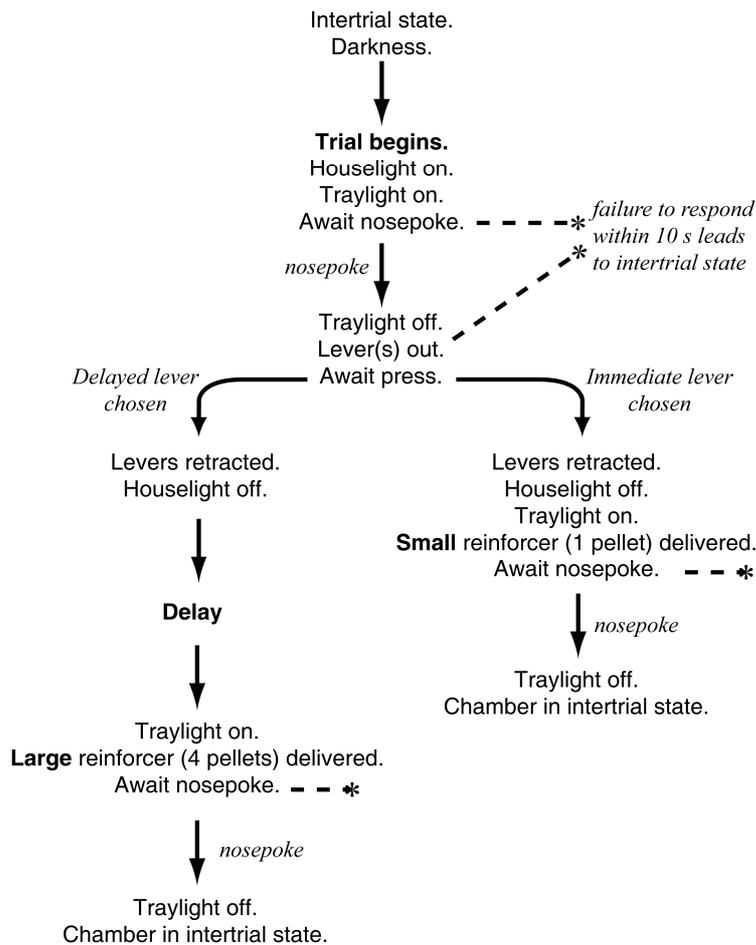
Figure 15: Task schematic: choice between small, immediate and large, delayed rewards

Hungry rats regularly choose between two levers. Responding on one lever leads to the immediate delivery of a small food reward (1 pellet); responding on the other leads to a much larger food reward (4 pellets), but this reward is delayed for between 0 and 60 seconds. The figure shows the format of a single trial; trials begin at regular intervals (every 100 s), so choice of the small reinforcer is always suboptimal. Sessions consist of 5 blocks. In each block, two single-lever trials are given (one trial for each lever), to ensure the animals sample the options available at that time; these are followed by ten choice trials. The delay to the large reinforcer is varied systematically across the session: delays for each block are 0, 10, 20, 40, and 60 s respectively. In the so-called "signalled" or "cue" condition, a stimulus light is illuminated during the delay to the large reinforcer; this is absent in the "unsignalled" or "no cue" condition, used for lesion studies. From Cardinal *et al.* (2000; 2001); based on a task by Evenden & Ryan (1996).
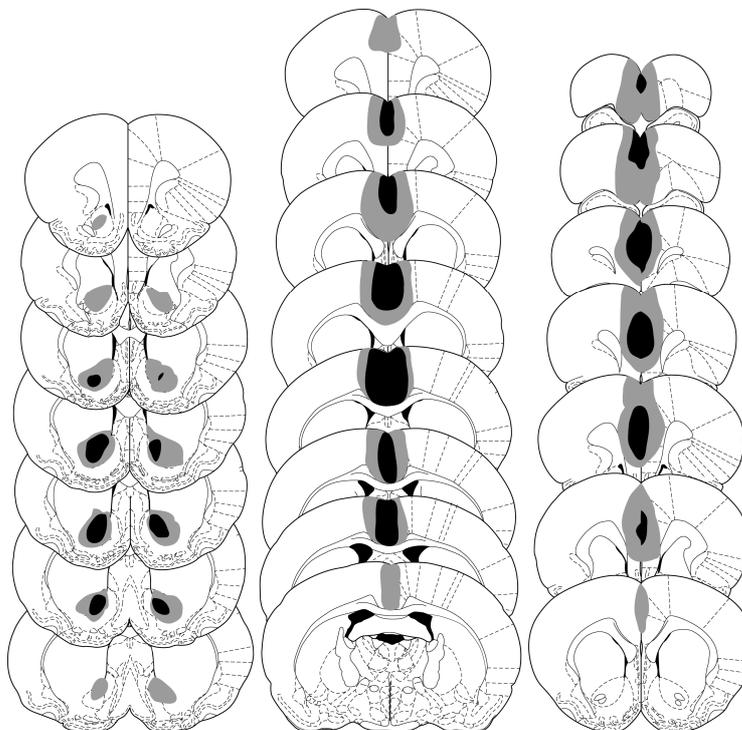
Figure 16: Schematics of lesions of the nucleus accumbens core (AcbC), anterior cingulate cortex (ACC), and medial prefrontal cortex (mPFC)

Schematics of lesions of the AcbC **(left)**, ACC **(middle)**, and mPFC **(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 +2.7 through +0.48 mm (AcbC), +2.7 mm through −1.3 mm (ACC), and +4.7 through +1.7 mm (mPFC) relative to bregma. Outlines are taken from Paxinos and Watson (1998). Figure from Cardinal *et al.* (2001).
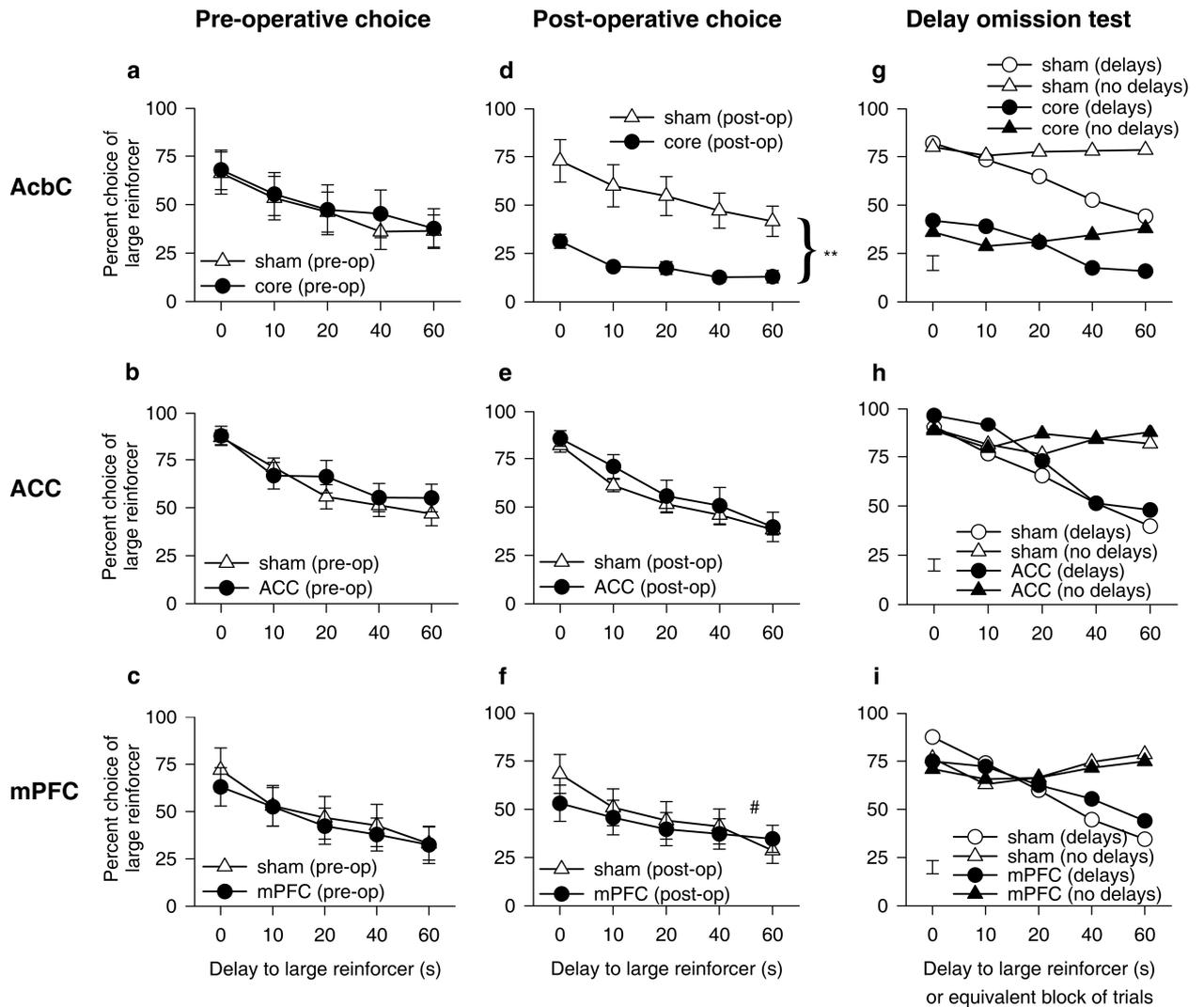
Figure 17: Choice between immediate, small and large, delayed rewards in rats with lesions of the AcbC, ACC, or mPFC

Effect of lesions of the AcbC (top), ACC (middle), or mPFC (bottom) on choice of delayed reward (● lesioned group; △ corresponding sham group; error bars, SEM). The "no cue" condition (see Figure 15) was used throughout. **Panels a–c** show 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 & Ryan, 1996; Cardinal *et al.*, 2000). **Panels d–f** illustrate choice in the first 7 postoperative 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). **Panels g–i** illustrate the effects of omitting all delays in alternating sessions (●/○, lesioned/sham groups with delays; ▲/△, 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. From Cardinal *et al.* (2001).
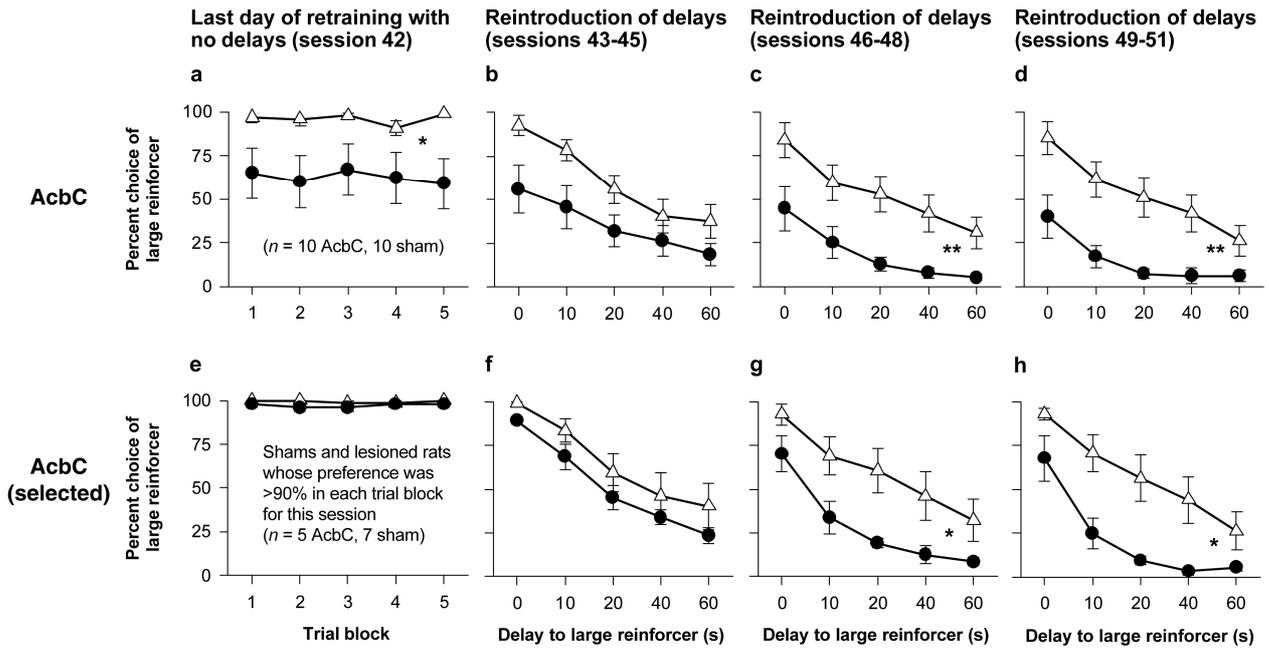
Figure 18: Further testing of AcbC-lesioned rats in the delayed reinforcement choice task
Panel a illustrates the preference of AcbC-lesioned rats following extended training in the absence of any delays (a further six sessions after completion of other behavioural tests; ● AcbC-lesioned group; Δ shams; error bars, SEM). **Panels 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. Panels 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. The subjects were the same as those reported in Cardinal *et al.* (2001); data from Cardinal *et al.* (2003b).

1.11.1.2 Processing of reward magnitude

Is the impulsive choice seen in AcbC-lesioned rats (Cardinal *et al.*, 2001) due to an effect on subjects' processing of reward delay, or of reward magnitude? As this task involves choice between reinforcers that differed in both magnitude and delay, impulsive choice might arise as a result of altered sensitivity to reinforcer magnitude, or delay, or both (Ho *et al.*, 1999) (Figure 19). Lesioned rats might have chosen the immediate small reward because they did not perceive the large reward to be as large (relative to the small reward) as sham-operated controls did, in which case the abnormally low magnitude of the large reward would be insufficient to offset the normal effects of the delay. Alternatively, they might have perceived the reward magnitudes normally, but been hypersensitive to the delay.
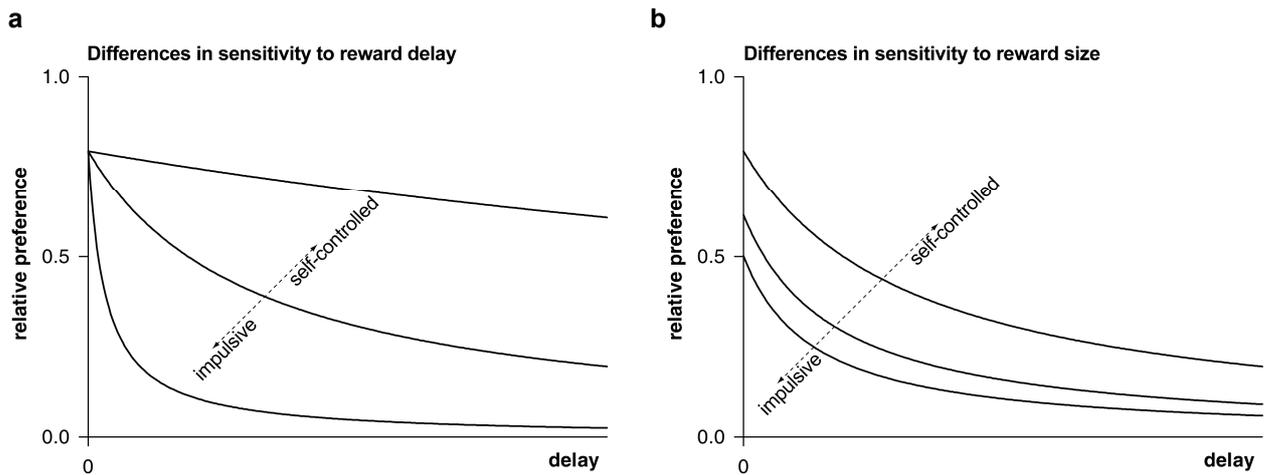
Figure 19: Delay and magnitude discounting applied to choice

Differences in delay discounting and differences in magnitude discounting can both affect choice between a smaller/sooner (SS) and a larger/later (LL) reward. In these theoretical curves, choice is calculated according to the multiplicative hyperbolic model of Ho *et al.* (1999). Subjects calculate value according to the formula

$$V = \frac{1}{1 + K \cdot d} \times \frac{V_{max}}{1 + Q/q}$$

where V is overall value, d is delay, q is quantity, V_{max} is the maximum possible value that a reward can have, K is a temporal (delay) discounting parameter and Q is a quantity (magnitude) discounting parameter. Subjects then allocate preference (e.g. relative number of choices or relative response rate) in proportion to the values of the two alternatives, i.e. $\dfrac{pref_A}{pref_A + pref_B} = \dfrac{V_A}{V_A + V_B}$. **(a)** These curves show the effect of altering K while holding Q constant. "Impulsive" subjects—those who are more likely to choose the SS reward over the LL reward—have a higher value of K than "self-controlled" subjects; their Q parameters do not differ. **(b)** These curves show the effect of altering Q while holding K constant. "Impulsive" subjects have a lower value of Q than "self-controlled" subjects; their K values do not differ.

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 18a), whereas in the model of Ho *et al.* (1999), relative preference between a one-pellet and a four-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.

However, hypersensitivity to the effects of delay appears somewhat more likely than alterations in reward magnitude processing as an explanation for the effects of AcbC lesions. As discussed above, AcbC lesions also reduced preference for the large reinforcer somewhat at zero delay (Cardinal *et al.*, 2001), but this was probably due to a task artefact, namely within-session generalization from trials in which delays were present (Cardinal *et al.*, 2000). When delays were consistently absent, AcbC-lesioned rats preferred the larger reward to the smaller (Cardinal *et al.*, 2001; 2003b). Further evidence supports the assertion that AcbC-lesioned rats can discriminate large from small rewards. Excitotoxic lesions of the whole Acb do not prevent rats from detecting changes in reward value, induced either by altering the concentration of a sucrose reward or by changing the deprivational state of the subject (Balleine & Killcross, 1994). Such lesions also do not impair rats' ability to respond faster when environmental cues predict the availability of larger rewards (Brown & Bowman, 1995), and nor does inactivation of the Acb with local anaesthetic

or blockade of α-amino-3-hydroxy-5-methyl-4-isoxazolpropionate (AMPA) glutamate receptors in the Acb (Giertler *et al.*, 2004); the effects of intra-Acb NMDA receptor antagonists have varied (Hauber *et al.*, 2000; Giertler *et al.*, 2003). AcbC-lesioned rats can still discriminate large from small rewards (Cardinal *et al.*, 2003b; 2004). Similarly, DA depletion of the Acb does not affect the ability to discriminate large from small reinforcers (Salamone *et al.*, 1994; Cousins *et al.*, 1996; Salamone *et al.*, 2001). However, these studies do not address the question of whether AcbC lesions alter the *quantitative* assessment of reward magnitude—e.g. whether such lesions alter Q in the model of Ho *et al.* (1999). Systemic DA antagonists do not affect the perceived quantity of food as assessed in a psychophysical procedure (Martin-Iverson *et al.*, 1987), but this is uninformative as to the role of the AcbC specifically.

The observation that AcbC lesions reduce preference for delayed, certain rewards (Pothuizen *et al.*, 2005) as well as delayed, large rewards (Cardinal *et al.*, 2001), is also consistent with the hypothesis that AcbC-lesioned animals have an impaired tolerance for delays, and that the effects are not due simply to effects on reward magnitude processing—though at present, the role of the AcbC in choosing uncertain reinforcement is also unclear (discussed further below). Acb lesions have also produced delay-dependent impairments in a delayed-matching-to-position task (Reading & Dunnett, 1991).

1.11.1.3 The matching law and reinforcer magnitude assessment

Semi-quantitative assessment of delay and magnitude discounting may be possible in delay-of-reinforcement choice tasks using indifference-point methodology (Ho *et al.*, 1999). Alternatively, relative preference for two reinforcers may be inferred from the distribution of responses on concurrent variable interval (VI) schedules of reinforcement. This literature stems from the discovery by Herrnstein (1961; 1970) of the "matching law". Herrnstein (1961) trained pigeons to respond on two concurrent VI schedules, and varied the relative availability of reinforcement on the two schedules while holding the overall reinforcement rate constant. He observed that the proportion of the total behaviour allocated to each response key approximately matched the proportion of reinforcers allocated to that key. This defines the matching law:

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

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

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

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 large body of work has sought to define the effects of varying parameters of reinforcement (such as rate, probability, delay, and magnitude) based on this principle (see de Villiers & Herrnstein, 1976).

This technique has not been without problems; in many circumstances, subjects have been found to "overmatch" (exhibit preferences that are exaggerated relative to the predictions of the matching law) or "undermatch" (exhibit reduced relative preferences). This has led to further development of the mathematical models (Baum, 1974; Baum, 1979), though it has been argued in some cases that this approach is circular (Rachlin, 1971). Maximum response rates (k in the equation above) have been shown to vary with the kind of reinforcement used (Belke, 1998), violating an assumption of Herrnstein's law. Nevertheless, the matching law and its extensions do a good job of describing the relationship between reinforcement rate and behaviour on concurrent VI and concurrent-chain schedules (Williams, 1994).

This allows for the possibility of assessing the effects of AcbC lesions upon reinforcer magnitude perception semi-quantitatively (Cardinal, 2001). Used with identical schedules delivering large and small rewards, the matching technique 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.

1.11.1.4 Learning with delayed reinforcement

If AcbC lesions do indeed induce hypersensitivity to delays of reinforcement, then the effects of AcbC lesions might also extend to *learning* with delayed reinforcement, as well as choice involving delayed reinforcers. In order to learn which actions are the correct ones that eventually lead to reinforcement, some mechanism must "bridge" the delay between action and outcome. Action–outcome delays impair instrumental learning in normal animals to some degree (Grice, 1948; Lattal & Gleeson, 1990; Dickinson *et al.*, 1992). If the AcbC is critical for learning with delayed reinforcement, then AcbC lesions should induce a delay-dependent impairment in free-operant learning with action–outcome delays. This prediction has not yet been tested.

1.11.1.5 Choice involving uncertain reward

Correlational studies have also suggested that the Acb may also be involved in the processing of uncertain or probabilistic reinforcement. DA neurons that innervate the Acb may fire in a manner related to reward probability (Fiorillo *et al.*, 2003; Fiorillo *et al.*, 2005; Niv *et al.*, 2005; Tobler *et al.*, 2005) and the midbrain, the site of the cell bodies of these neurons, responds to stimulus uncertainty in humans (Aron *et al.*, 2004). A greater blood-oxygen-level-dependent (BOLD) response is observed in the human Acb during the selection of high-reward/high-risk options, compared to low-reward/low-risk outcomes, in a task where the risk is of not winning (Ernst *et al.*, 2004), with similar activation to high-reward/high-risk option selection in a task where the risk is of losing (Matthews *et al.*, 2004); this latter activation was correlated with personality measures of harm avoidance. Likewise, an increase in Acb activation (BOLD signal) preceded risk-taking decisions in a financial game with human subjects (Kuhnen & Knutson, 2005). However, to date no studies have examined the contribution of the AcbC to choice involving reward uncertainty. In a recent interventional study, AcbC lesions reduced preference for delayed, certain rewards (Pothuizen *et al.*, 2005), but this does not specifically address the contribution of the AcbC to choosing rewards based on their certainty, particularly as there is evidence suggesting that AcbC lesions impair the processing of delayed reinforcement (Cardinal *et al.*, 2001; Pothuizen *et al.*, 2005).

1.11.1.6 Relationship to neuromodulator function

The Acb is innervated by the a number of neuromodulator systems, including 5-HT (see Halliday *et al.*, 1995) and DA (Ungerstedt, 1971; Fallon & Loughlin, 1995). The DA projection to the Acb is prominent, but although systemic D_2-type DA receptor antagonists can induce impulsive choice involving delayed reinforcement (Wade *et al.*, 2000), this effect may not depend critically on DA in the Acb. Intra-Acb D_1 and D_2 receptor antagonists do not affect rats' ability to wait for reward in a cued progressive delay task (Wakabayashi *et al.*, 2004), and DA depletion of the Acb using 6-OHDA appears not to affect delay discounting directly, though it modifies the effect of systemic 5-HT_{1A} receptor agonists on choice between SS and LL rewards (Winstanley *et al.*, 2005b). The Acb does not receive a substantial NA innervation (Aston-Jones *et al.*, 1995).

1.11.2 Nucleus accumbens shell (AcbSh)

In contrast to the effects of AcbC lesions on choice between delayed, certain and immediate, uncertain rewards, AcbSh lesions have been shown to have effects neither on this task nor on DRL efficiency (Pothuizen *et al.*, 2005). The AcbSh responds to a variety of USs (Bassareo & Di Chiara, 1999; Ito *et al.*, 2000) and has a role in the hedonic assessment of rewards (Berridge, 2000; Kelley & Berridge, 2002). It plays a role in latent inhibition (Pothuizen *et al.*, 2005; 2006), and influences unlearned behaviours including feeding (Kelley & Swanson, 1997; Stratford & Kelley, 1997; Basso & Kelley, 1999; Kelley, 1999) and locomotion (Swanson *et al.*, 1997; Parkinson *et al.*, 1999a). The AcbSh has also been shown to be abnormal in animal models of ADHD (Papa *et al.*, 1996; Carey *et al.*, 1998; Papa *et al.*, 1998; Sadile, 2000). However, these results suggest it does not contribute to choice involving delayed or uncertain rewards (Pothuizen *et al.*, 2005).

1.11.3 Anterior cingulate cortex (ACC)

Excitotoxic lesions of the ACC (Figure 16) have no effect on choice between SS and LL rewards in rats (Cardinal *et al.*, 2001) (Figure 17), indicating that the ACC is not required for normal choice of delayed reinforcement. These results suggest that ACC dysfunction is not an important contributor to impulsive choice, despite the involvement of the ACC in reward-related learning (Bussey *et al.*, 1997a; Bussey *et al.*, 1997b; Parkinson *et al.*, 2000c; Cardinal *et al.*, 2003a) and findings of ACC abnormalities in ADHD (Bush *et al.*, 1999; Rubia *et al.*, 1999). However, ACC lesions do impair choice between small/sooner/low-effort and large/later/high-effort alternatives, reducing preference for the high-effort option (Walton *et al.*, 2002; 2003), indicating that the ACC is involved in promoting the selection of effortful alternatives. The DA innervation of the ACC does not appear important for this function (Walton *et al.*, 2005).

However, ACC lesions can make rats motorically impulsive, with simple disinhibition or "execution" impulsivity. ACC-lesioned rats have been found to over-respond to unrewarded stimuli (Bussey *et al.*, 1997a; Parkinson *et al.*, 2000c) and to respond prematurely in situations where they are required to wait (Muir *et al.*, 1996). They also exhibit discriminative deficits in Pavlovian conditioning tasks (Cardinal *et al.*, 2003a), though the full range of functions associated with the ACC (including error detection, attentional control, and mood; see e.g. Devinsky *et al.*, 1995; Volkow *et al.*, 1997; Maas *et al.*, 1998; Childress *et al.*, 1999; Bush *et al.*, 2000; Garavan *et al.*, 2000; Paus, 2001; Cardinal *et al.*, 2002a; Vogt, 2005) is beyond the scope of this thesis.

The contribution of the ACC to probabilistic choice is less clear. In both humans and rhesus monkeys, the ACC responds to anticipated gain in tasks in which rewards of different magnitudes are available with

varying probabilities. In the rhesus monkey, the ACC responds to some combination of reward size and reward probability (Amiez *et al.*, 2005a) and deactivation of the ACC impairs such choices (Amiez *et al.*, 2005b), but human studies would suggest that the ACC responds to the magnitude rather than the probability of expected gains (Rogers *et al.*, 2004b). However, a nearby region of human medial PFC has been observed to respond to reward probability (Knutson *et al.*, 2005).

1.11.4 Prelimbic (PrL) and infralimbic (IL) cortex

The mPFC projects to the AcbC, is involved in reward-related learning (e.g. Balleine & Dickinson, 1998; Richardson & Gratton, 1998; Bechara *et al.*, 1999; Tzschentke, 2000), receives DA and 5-HT input (see Fallon & Loughlin, 1995; Halliday *et al.*, 1995), and has been observed to be abnormal in ADHD (Ernst *et al.*, 1998; Rubia *et al.*, 1999). However, lesions of the rat mPFC, primarily PrL and infralimbic (IL) cortex (Figure 16), had no delay-specific effects on choice between large/delayed and small/immediate rewards (Cardinal *et al.*, 2001) (Figure 17); the effects observed appeared to be task-specific, related to an insensitivity to the contingencies or stimuli present in the task, perhaps as a result of a loss of temporal discriminative stimulus control (Cardinal *et al.*, 2003b). It is important to note that PrL may have more functional homology to the primate dorsolateral PFC than to regions that are medial within human PFC (Uylings *et al.*, 2003). Aspirative lesions of the mPFC have previously been shown to induce a deficit in timing ability in rats (Dietrich & Allen, 1998), with impaired temporal discrimination in the peak procedure, an operant task that assesses the ability to time a stimulus (Catania, 1970; Roberts, 1981). Consistent with the view that mPFC lesions did not affect the basic process of choosing between reinforcers of different value, combined PrL/IL lesions did not affect choice between small/low-effort and large/high-effort alternatives in the task of Walton *et al.* (2003).

1.11.5 Orbitofrontal cortex (OFC)

The OFC is a region of the PFC that projects to the AcbC and is strongly implicated in the assessment of reward value. Mobini *et al.* (2002) recently found that lesions encompassing the OFC induced impulsive choice in a discrete-trial SS/LL reward choice task very similar to that described above (Figure 20). As before, results from this simple form of task do not indicate whether the impulsive choice was as a result of altered sensitivity to reinforcer magnitude or delay. Although these lesions damaged prelimbic cortex (PrL) in addition to the OFC (Mobini *et al.*, 2002), the hypothesis that OFC damage was responsible for the behavioural effect is strengthened by the finding that mPFC lesions encompassing PrL do not induce impulsive choice (Cardinal *et al.*, 2001). In contrast, Winstanley *et al.* (2004b) recently found that OFC lesions induced the opposite effect—better self-control than shams (Figure 21)—in exactly the task described above (Figure 15, p. 38). One possible reason for this discrepancy is that subjects in Winstanley *et al.*'s study were trained before the OFC was destroyed and retested postoperatively, while Mobini *et al.* trained and tested postoperatively. Another is that Mobini *et al.* (2002) offered rats a choice between a one-pellet immediate reinforcer and a two-pellet delayed reinforcer, whereas Winstanley *et al.* (2004b) used a one-pellet immediate reinforcer and a four-pellet delayed reinforcer. As discussed above, differences in subjects' sensitivity to either the delay or the magnitude of reinforcement can play a role in determining preference in this task (Ho *et al.*, 1999; Mobini *et al.*, 2002; Cardinal *et al.*, 2003b) and it may be that OFC lesions affect both, increasing both the delay discounting parameter K and the magnitude discounting parameter Q (Mobini *et al.*, 2002). An increase in K would imply steeper delay discounting; an increase in Q would imply an increase in sensitivity to the ratio of the magnitudes of the two reinforcers, and could mask (or potentially reverse) the increase in impulsivity produced by the increase in K.
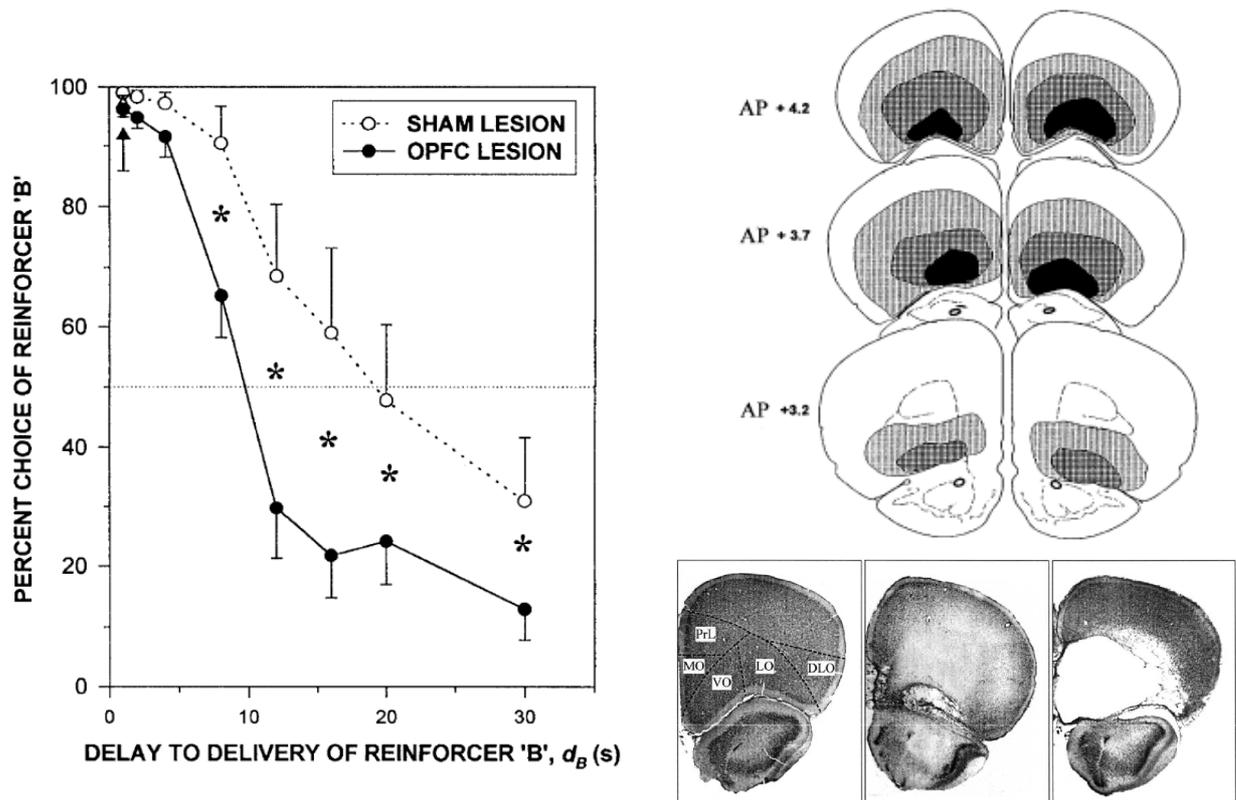
Figure 20: Choice between immediate, small and large, delayed rewards in rats with lesions of the orbitofrontal cortex (OFC)

In this study, rats received excitotoxic lesions of the OFC or sham lesions before being trained on a delayed reinforcement choice task similar to that described earlier, but involving choice between a one-pellet small reward and a two-pellet large reward. Delays to the large reinforcer varied across, rather than within, sessions. **(Left)** OFC-lesioned rats were more impulsive than shams, choosing the larger/later reward less often. **(Right)** Lesion schematics and representative photomicrographs. From Mobini *et al.* (2002).

There is direct support for this hypothesis: OFC lesions appear to increase *K*, the rate of delay discounting, as well as increasing the magnitude sensitivity parameter *Q* (Kheramin *et al.*, 2002; Kheramin *et al.*, 2003). The same effect of increases in both *K* and *Q* has been observed with DA-depleting OFC lesions (Kheramin *et al.*, 2004). This emphasizes the necessity for quantitative analysis of delay and magnitude sensitivity (Ho *et al.*, 1999) or the use of multiple, very different paradigms to provide independent measurements of sensitivity to delay and magnitude (Cardinal *et al.*, 2003b). It also reminds us of an important clinical point: faced with steep delay discounting in a task involving choice between SS and LL rewards, increasing the ratio of the large to the small reward may ameliorate the impulsivity.

As discussed above, it has been suggested that hyperbolic discounting is explicable as the overall effect of two or more different systems, such as an explicit (declarative) system that exhibits minimal or exponential discounting, plus phenomena that make rewards more salient and promote their choice when they are immediately available. Recently, such a two-factor model was used in the analysis of a functional magnetic resonance imaging (fMRI) study of choice involving rewards differing in magnitude and delays, with delays ranging from less than a day to 6 weeks (McClure *et al.*, 2004): lateral prefrontal and intraparietal cortical regions were activated independently of the delay, and were suggested to be part of a system that evaluates both immediate and delayed rewards according to a "rational" (meaning non-hyperbolic) temporal discounting system, while limbic regions including the ventral striatum and medial OFC were preferentially activated by the relatively immediate rewards, and were suggested to be part of a
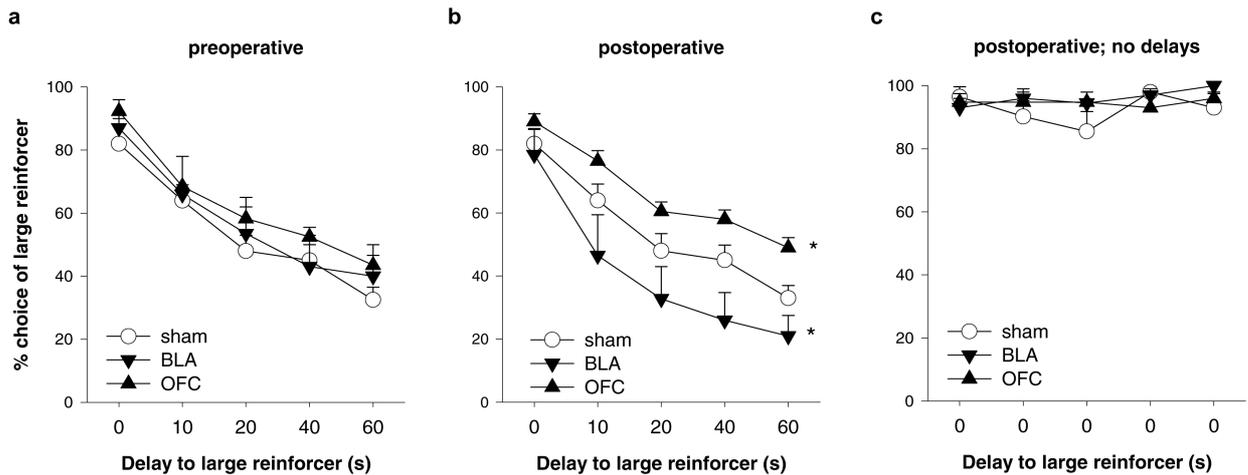
Figure 21: Choice between immediate, small and large, delayed rewards in rats with lesions of the basolateral amygdala (BLA) or OFC

Rats were trained on the delayed reinforcement choice task (involving choice between a one-pellet immediate reinforcer and a four-pellet large reinforcer, and within-session changes in the delay to the large reinforcer; Figure 15) before matched groups of subjects **(a)** received excitotoxic lesions of the BLA, or the OFC, or sham lesions; **(b)** shows their postoperative performance, in which BLA-lesioned subjects were more impulsive (more likely to choose a single immediate pellet over a four-pellet delayed reward) and OFC-lesioned subjects were less impulsive than sham-operated controls (* $p < .05$, difference from shams). **(c)** When all delays were removed from the task, BLA- and OFC-lesioned subjects chose identically to shams, preferring the four-pellet reward to the one-pellet reward. Redrawn from Winstanley *et al.* (2004b).

system that promotes the choice of imminent rewards without consideration of delayed alternatives. These limbic regions were more likely to be activated than the "delay-independent" areas on trials in which an earlier reward was chosen. This would sit neatly with studies showing that OFC lesions reduce impulsive choice (Winstanley *et al.*, 2004b); however, it does not square so easily with rodent evidence showing that destruction of the AcbC (a major part of the ventral striatum) or the OFC enhances delay discounting, meaning that delayed alternatives are less likely to be chosen (Cardinal *et al.*, 2001; Kheramin *et al.*, 2002; Mobini *et al.*, 2002; Kheramin *et al.*, 2003).

The PFC, which projects heavily to the AcbC (Brog *et al.*, 1993), is also involved in decision making under conditions of uncertainty. Humans with OFC or ventromedial PFC damage are impaired in the Iowa gambling task (Bechara *et al.*, 1994; 1996; 1997), in which subjects must learn to differentiate between low-reward, low-risk card decks that yield a net positive outcome and high-reward, high-risk decks that yield a net negative outcome, though the precise locus and nature of the deficit seen in this task is debated (Manes *et al.*, 2002; Clark *et al.*, 2003; Fellows & Farah, 2005). OFC neurons respond to reward expectancy (see Hikosaka & Watanabe, 2000). Choice between small, likely rewards and large, unlikely rewards increases blood flow and BOLD signal in orbital and inferior PFC (Rogers *et al.*, 1999b; Ernst *et al.*, 2004; Rogers *et al.*, 2004b), and OFC damage also impairs performance of a task requiring human subjects to choose between two possible outcomes and to bet on their choice, with lesioned subjects deciding slowly and failing to choose the optimal, most likely outcome (Rogers *et al.*, 1999a). Excitotoxic lesions of the OFC make rats less likely than sham-operated controls to choose a large, uncertain reward over a small, certain reward (Mobini *et al.*, 2002); OFC-lesioned rats had lower indifference odds (higher indifference probabilities; steeper uncertainty discounting) and exhibited risk-averse choice. As discussed above, there is direct evidence that excitotoxic OFC lesions and OFC DA depletion do alter sensitivity to the relative magnitudes of the two rewards (Kheramin *et al.*, 2004; Kheramin *et al.*, 2005), but the effect

of steepening uncertainty discounting (of increasing the odds discounting parameter *H*) is present in addition to the effects on reinforcer magnitude sensitivity (Kheramin *et al.*, 2003).

A recent fMRI study also examined human regional cerebral BOLD responses to decision making in a task that explicitly distinguished "ambiguity" from "risk" (Hsu *et al.*, 2005). "Ambiguity" referred to the situation in which the probability of a successful outcome was unknown—such as having to bet on whether the next card drawn from a twenty-card deck containing red or blue cards would be red or blue, with no further information. "Risk" referred to the situation in which the probability of a successful outcome was known, but not zero or one—such as having to bet on the colour of the next card drawn from a deck containing ten red and ten blue cards. This task produces behaviour that is economically self-contradictory. Subjects prefer to bet on red from the risky deck than on red from the ambiguous deck, but also prefer to bet on blue from the risky deck than on blue from the ambiguous deck (Becker & Brownson, 1964; MacCrimmon, 1968). These preferences are mutually inconsistent under simple probability theory—termed the Ellsberg paradox (Ellsberg, 1961)—and imply an inherent aversion to ambiguity. In these tasks, the OFC, amygdala, and dorsomedial PFC were more active under conditions of ambiguity than risk, while dorsal striatal activity showed the opposite pattern (risk > ambiguity). While normal humans exhibited aversion to ambiguity, and also aversion to risk (in other tasks in which the card proportion was varied), humans with OFC lesions were averse neither to risk nor ambiguity—behaviourally abnormal, but consistent with expected utility theory (Hsu *et al.*, 2005).

1.11.6 Insula

A further cortical region that may be involved in decisions involving uncertainty is the insula, or insular cortex. Anterior insula activation has been observed to precede risk-averse choice in humans (Kuhnen & Knutson, 2005), in a task in which Acb activation preceded risk-prone choice. The authors suggest that in tasks such as these, the Acb represents predictions of gain (Knutson *et al.*, 2001), while the insula represents predictions of loss (see also Paulus *et al.*, 2003); activation in both structures is related to personality measures of harm avoidance (Paulus *et al.*, 2003; Matthews *et al.*, 2004).

1.11.7 Basolateral amygdala (BLA)

The basolateral amygdala (BLA) also projects to the AcbC, and has extensive reciprocal connections with the OFC. Excitotoxic lesions of the BLA promote impulsive choice in a task involving choice between an immediate one-pellet reward and a delayed four-pellet reward (Winstanley *et al.*, 2004b) (Figure 21), similar to the effects of AcbC lesions (Cardinal *et al.*, 2001) but opposite to those of OFC lesions in the same task (Winstanley *et al.*, 2004b) (Figure 21). Although this study is notable for finding opposite effects of BLA and OFC lesions, which is unusual (see also Izquierdo & Murray, 2005), the explanation for this effect is unclear. One obvious possibility, given the effects of OFC lesions to increase both the delay discounting parameter *K* and the magnitude sensitivity parameter *Q* (in the model of Ho *et al.*, 1999), is that BLA lesions and AcbC lesions simply increase *K* without affecting *Q* (cf. Figure 19, p. 41). There is indirect evidence for this in the case of AcbC lesions, discussed above; for the BLA, this hypothesis remains untested. Some studies have demonstrated deficits following amygdala inactivation when reward size is suddenly changed (Salinas *et al.*, 1993; Coleman-Mesches *et al.*, 1996; Salinas & McGaugh, 1996; Salinas *et al.*, 1997; Liao & Chuang, 2003), though changing the size of a reward for performing the same task has obvious emotional significance and the amygdala is well known to be involved in affective representation (see Aggleton, 2000; Cardinal *et al.*, 2002a). One study has found deficits in *memory* for reinforcer magnitude following amygdala lesions, even if this was not a primary

deficit in reinforcer magnitude discrimination (Kesner & Williams, 1995). None of these bear directly on the question of whether relative reinforcer magnitude discrimination (as measured by Q) is altered by BLA lesions.

A recent study has also suggested the involvement of the BLA in promoting the selection of effortful alternatives. Floresco & Ghods-Sharifi (2006) showed that BLA inactivation with the local anaesthetic bupivacaine impaired rats' ability to choose a large/high-effort alternative over a small/low-effort alternative. This is much like the effect of ACC lesions discussed above (Walton *et al.*, 2002; 2003), and indeed, a reversible BLA–ACC disconnection lesion also impaired selection of large/high-effort alternatives (Floresco & Ghods-Sharifi, 2006), suggesting that direct information transfer between the BLA and the ACC is important in this task.

1.11.8 Subthalamic nucleus (STN)

The subthalamic nucleus (STN) is a component of the basal ganglia that receives projections both from the globus pallidus (pallidum) and the cerebral cortex (Alexander & Crutcher, 1990; Hamani *et al.*, 2004) and projects to basal ganglia relay structures (including the globus pallidus, the rodent homologue of the external part of the primate globus pallidus) and output structures of the basal ganglia, including the entopeduncular nucleus and the substantia nigra pars reticulata (Heimer *et al.*, 1995; Hamani *et al.*, 2004), which project on to thalamus and thence to cortex. Lesions of the STN decreased impulsive choice in a task involving choice of a single immediate food pellet or four pellets delivered after a delay (Winstanley *et al.*, 2005a), a task in which OFC lesions had the same effect (Winstanley *et al.*, 2004b). STN lesions also impaired autoshaping (Winstanley *et al.*, 2005a), meaning locomotor approach to appetitive Pavlovian CSs (Brown & Jenkins, 1968; Williams & Williams, 1969). However, this is unlikely to explain the effect of STN lesions to promote choice of LL rewards—not least because AcbC lesions also impair autoshaping (Parkinson *et al.*, 2000c; Cardinal *et al.*, 2002b) but reduce choice of LL rewards (Cardinal *et al.*, 2001), while ACC lesions impair autoshaping (Bussey *et al.*, 1997a; Parkinson *et al.*, 2000c; Cardinal *et al.*, 2003a) but do not affect choice between SS/LL rewards (Cardinal *et al.*, 2001), but more simply because there was no explicit CS in this task differentially associated with the two rewards, and approach to which would promote choice of the SS reward. Furthermore, STN lesions tend to increase premature responding, often thought of as an index of motor impulsivity (Baunez & Robbins, 1997; Baunez *et al.*, 2001). It is not known whether STN lesions affect reward magnitude discrimination or uncertainty discounting.

1.11.9 Hippocampus (H)

Finally, a role of the hippocampus in learning with delayed reinforcement might be suspected. As discussed earlier, contextual conditioning is important in learning with delays, and there is good evidence that the hippocampus contributes to the representation of context. Context-specific representations develop in the hippocampus (Smith & Mizumori, 2006), and lesions of the hippocampal formation (H) have been shown to impair Pavlovian conditioning to a contextual CS, but not to a discrete CS, in rats (Hirsh, 1974; Selden *et al.*, 1991; Kim & Fanselow, 1992; Phillips & LeDoux, 1992; Honey & Good, 1993; Jarrard, 1993; Kim *et al.*, 1993; Phillips & LeDoux, 1994; Phillips & LeDoux, 1995; Chen *et al.*, 1996; Maren & Fanselow, 1997; Anagnostaras *et al.*, 1999; Rudy *et al.*, 2002), at least for some processes involving contextual representation (Good & Honey, 1991; Holland & Bouton, 1999; Good, 2002). In some cases, discrete CS conditioning has even been enhanced (e.g. Ito *et al.*, 2005). Since context–outcome associations are thought to hinder instrumental learning with delayed reinforcement through contextual

competition (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994), it follows that if H lesions impair the formation of associations involving the context, such lesions might reduce contextual competition and hence *facilitate* instrumental conditioning when there is an action–outcome delay.

Despite this clear prediction, the contribution of the hippocampus to learning with delayed reinforcement, or to self-controlled choice, has not previously been investigated in detail. One previous study found that aspirative lesions of the dorsal hippocampus did not affect appetitive instrumental conditioning with delayed reinforcement (Port *et al.*, 1993), but this study was poorly designed to address this question in a number of ways; amongst its flaws, the study used aspirative rather than excitotoxic lesions, used a task in which alterations in response rates affected the instrumental contingency, and tested subject at a single delay with no zero-delay control condition.

The only study to date to address the influence of the hippocampus on choice involving delayed or uncertain reward was that of Rawlins *et al.* (1985), who examined choice between certain and uncertain rewards. Normal rats preferred immediate certain reward to immediate uncertain reward, and also preferred delayed certain reward to immediate uncertain reward; however, rats with hippocampal or medial septal lesions were less tolerant of the delay (or more tolerant of the uncertainty), preferring immediate uncertain reward to delayed certain reward. However, this study does not answer the question of whether hippocampal lesions affect the processing of reward delay or reward uncertainty specifically.

1.12 OUTLINE OF EXPERIMENTAL WORK IN THIS THESIS

This thesis has three principal objectives: first, to establish whether the role of the AcbC in choosing large, delayed rewards reflects an underlying deficit in the processing of reward delay or of reward magnitude; second, to investigate the role of the AcbC in decisions involving uncertain reward; and third, to establish the role of the hippocampus in the processing of delayed reward. Chapter 2 will examine the role of the AcbC in free-operant learning with delayed reward, performance of a previously learned free-operant response for delayed reward, and the quantitative allocation of behaviour to match obtained reward magnitudes. Chapter 3 will examine the role of the hippocampus in learning with delayed reward, performance of a previously learned free-operant response for delayed reward, and choice between SS and LL reward alternatives. Chapter 4 will return to the AcbC, examining its role in choice between small/certain and large/unlikely rewards.

Chapter 2: Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat, but do not impair reinforcer magnitude discrimination

2.1 ABSTRACT

Background: Delays between actions and their outcomes severely hinder reinforcement learning systems, but little is known of the neural mechanism by which animals overcome this problem and bridge such delays. The nucleus accumbens core (AcbC), part of the ventral striatum, is required for normal preference for a large, delayed reward over a small, immediate reward (self-controlled choice) in rats, but the reason for this is unclear. These experiments investigated the role of the AcbC in learning a free-operant instrumental response using delayed reinforcement, performance of a previously learned response for delayed reinforcement, and assessment of the relative magnitudes of two different rewards.

Results: Groups of rats with excitotoxic or sham lesions of the AcbC acquired an instrumental response with different delays (0, 10, or 20 s) between the lever-press response and reinforcer delivery. A second (inactive) lever was also present, but responding on it was never reinforced. As expected, the delays retarded learning in normal rats. AcbC lesions did not hinder learning in the absence of delays, but AcbC-lesioned rats were impaired in learning when there was a delay, relative to sham-operated controls. All groups eventually acquired the response and discriminated the active lever from the inactive lever to some degree. Rats were subsequently trained to discriminate reinforcers of different magnitudes. AcbC-lesioned rats were more sensitive to differences in reinforcer magnitude than sham-operated controls, suggesting that the deficit in self-controlled choice previously observed in such rats was a consequence of reduced preference for delayed rewards relative to immediate rewards, not of reduced preference for large rewards relative to small rewards. AcbC lesions also impaired the performance of a previously learned instrumental response in a delay-dependent fashion.

Conclusions: These results demonstrate that the AcbC contributes to instrumental learning and performance by bridging delays between subjects' actions and the ensuing outcomes that reinforce behaviour.

2.2 BACKGROUND

Animals learn to control their environment through instrumental (operant) conditioning. When an animal acts to obtain reward or reinforcement, there is often a delay between its action and the outcome; thus, animals must learn instrumental action–outcome contingencies using delayed reinforcement. Although such delays impair learning, animals can nevertheless bridge substantial delays to acquire instrumental responses (Dickinson *et al.*, 1992). Little is known of the neural basis of this process. However, abnormalities in learning from delayed reinforcement may be of considerable clinical significance (Rahman *et al.*, 2001). Impulsivity is part of the syndrome of many psychiatric disorders, including mania, drug addiction, antisocial personality disorder, borderline personality disorder, and ADHD (APA, 2000). Impulsive choice, one aspect of impulsivity (Evenden, 1999b), is exemplified by the tendency to choose small

rewards that are available immediately instead of larger rewards that are only available after a delay (Ainslie, 1975; 2001), and may reflect dysfunction of reinforcement learning systems mediating the effects of delayed rewards (Ainslie, 1975; Sagvolden & Sergeant, 1998).

The nucleus accumbens (Acb) responds to anticipated rewards in humans, other primates, and rats (Schultz *et al.*, 1992; Miyazaki *et al.*, 1998; Martin & Ono, 2000; Schultz *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; Cromwell & Schultz, 2003; Bjork *et al.*, 2004), and is innervated by dopamine (DA) neurons that respond to errors in reward prediction in a manner appropriate for a teaching signal (Schultz *et al.*, 1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Schultz, 2006). The Acb may therefore represent a reinforcement learning system specialized for learning with delayed reinforcement (Houk *et al.*, 1995; Wickens & Kötter, 1995). If this is the case, then damage to the Acb should not interfere with reinforcement learning in all circumstances, but should produce selective impairments in learning when reinforcement is delayed. This prediction has not previously been tested. However, lesions of the AcbC cause rats to prefer small immediate rewards (a single food pellet delivered immediately) to large delayed rewards (four pellets delivered after a delay); that is, AcbC-lesioned rats exhibit impulsive choice (Cardinal *et al.*, 2001; 2002a). The reason for this is not clear. It might be that AcbC-lesioned rats exhibit steeper temporal discounting, such that the subjective utility (value) of future rewards declines more rapidly than normal as the reward is progressively delayed (Cardinal *et al.*, 2003b; 2004). It might also be that AcbC-lesioned rats are less good at representing the contingency between actions and their outcomes when the outcomes are delayed, so that they choose impulsively because they are less certain or less aware that their choosing the delayed reward does in fact lead to that reward being delivered (Cardinal *et al.*, 2003b; 2004). Both explanations would reflect a problem in dealing with delayed reinforcement in AcbC-lesioned rats. However, there might be a simpler explanation for the impulsive choice exhibited by AcbC-lesioned rats: they might perceive the size (magnitude) of rewards differently. For example, if they do not perceive the delayed reward to be as large, relative to the immediate reward, as normal rats did, then they might choose impulsively despite processing the delays to reward normally, simply because the delayed reinforcer is not subjectively large enough to compensate for the normal effects of the delay (Ho *et al.*, 1999; Cardinal *et al.*, 2003b; 2004).

To investigate whether the AcbC is a reinforcement learning system specialized for delayed reinforcement, the ability of AcbC-lesioned rats to detect instrumental contingencies across a delay was first determined. The ability of AcbC-lesioned rats to acquire instrumental responding with delayed reinforcement was compared to that of sham-operated controls; each subject was allowed to respond freely on two levers, one of which produced reinforcement after a delay of 0, 10, or 20 s (Figure 22). AcbC lesions only retarded instrumental learning when reinforcement was delayed, demonstrating a role for the AcbC in bridging action–outcome delays during learning. Subsequently, to establish whether AcbC-lesioned rats perceive reward magnitude abnormally, these subjects' sensitivity to reinforcer magnitude was assessed by measuring their relative preference for two different reinforcers using concurrent interval schedules of reinforcement. Reinforcer magnitude discrimination in AcbC-lesioned rats in this task was at least as good as in sham-operated controls, consistent with previous evidence of reinforcer magnitude discrimination following lesions of the whole Acb (e.g. Balleine & Killcross, 1994; Brown & Bowman, 1995). Together, these results suggest that the impulsive choice seen in AcbC-lesioned rats (Cardinal *et al.*, 2001) is due to a problem in processing delayed reward, not in processing the magnitudes of the reward alternatives. Finally, to establish whether the AcbC is required for the performance of an instrumental response for delayed reinforcement, as well as for the learning of such a response, naïve rats were trained to respond for delayed reinforcement (Figure 22) before destroying the AcbC. Such lesions also

impaired performance of a previously learned instrumental response only when reinforcement was delayed, indicating that the AcbC makes an enduring contribution to bridging delays between subjects' actions and the ensuing outcomes.
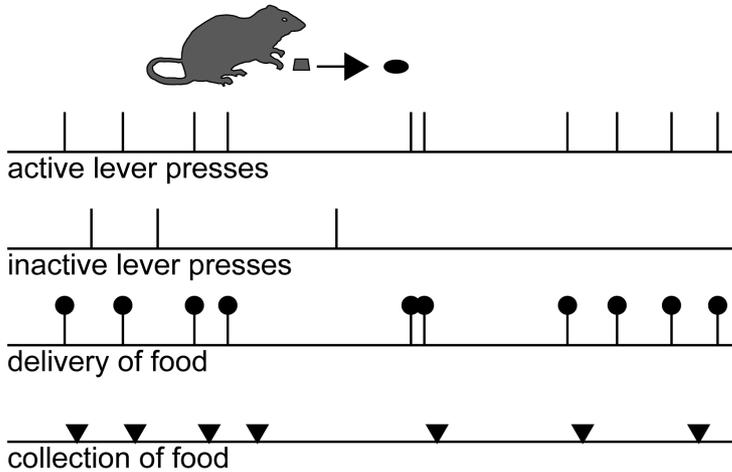
a) Zero delay
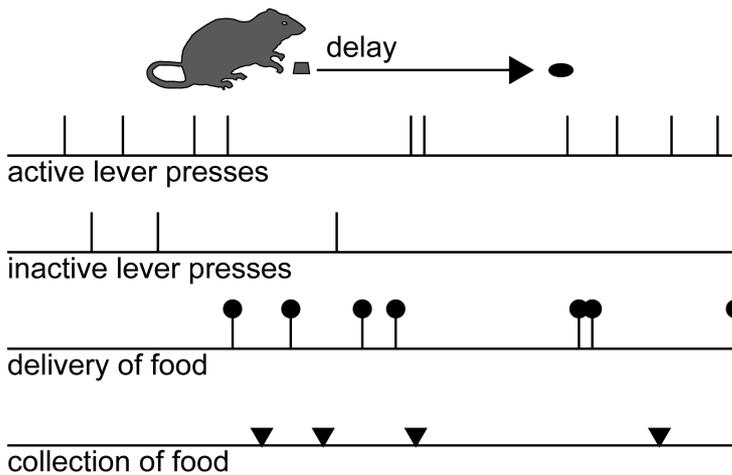
Figure 22: Task schematic: free-operant instrumental responding on an FR-1 schedule with delayed reinforcement

Subjects are offered two levers; one (the active lever) delivers a single food pellet for every press (an FR-1 schedule) and the other (the inactive lever) has no programmed consequence. Food can either be delivered **(a)** immediately or **(b)** after a delay following responses on the active lever. The levers remain available throughout the session (hence, free-operant responding: animals are free to perform the operant at any time). Events of interest are lever presses, delivery of food pellets, and collection of food by the rat (when it pokes its nose into the food alcove following food delivery). To obtain food, the hungry rat must discriminate the active from the inactive lever, which is more difficult when the outcome is delayed. In these examples, the rat's response patterns (active and inactive lever presses, and collection of food) are fictional, while food delivery is contingent upon active lever pressing.

2.3 METHODS

2.3.1 Overview of experiments

2.3.1.1 Experiment 1A: Effects of AcbC lesions on acquisition of instrumental responding with delayed reinforcement

Fifty naïve rats received excitotoxic lesions of the AcbC ($n = 26$) or sham lesions ($n = 24$). Two died postoperatively. Subjects were next trained in a task in which they had continuous access to two identical levers; one lever delivered a single food pellet each time it was pressed, and the other lever had no effect. For some rats, the food pellet was delivered immediately after the lever press (0 s condition; $n = 8$ AcbC-lesioned rats and 8 shams). For others, each pellet was delayed by either 10 s (8 AcbC, 8 sham) or 20 s (8 AcbC, 8 sham). Subjects were trained for 14 sessions.

2.3.1.2 Experiment 1B: Effects of AcbC lesions on the ability to match response distribution to reinforcer magnitude distribution

After the same rats had their locomotor activity assessed, they moved on to a task testing their ability to judge differences in the magnitude of two reinforcers. They were again offered two levers, but this time both levers delivered reinforcement on a VI schedule, which provides reinforcement in an intermittent and temporally unpredictable fashion. Reinforcers consisted of either 1 or 4 sucrose pellets. Over sessions, the levers' roles changed so that the ratio of the sizes of the reinforcers available on the two levers was 4:1, 1:1, or 1:4. Subjects' responding was measured to establish their ability to judge the relative differences in reinforcer magnitudes and to allocate their responses according to the matching law (Herrnstein, 1961; Herrnstein, 1970; Williams, 1994). Finally, they were killed and perfused for histology.

2.3.1.3 Experiment 2: Effects of AcbC lesions on performance of a previously learned instrumental response for delayed reinforcement

A further 48 naïve rats were trained to acquire an instrumental response as before, with delays to reinforcement of 0 s ($n = 16$), 10 s ($n = 16$), or 20 s ($n = 16$). One rat spontaneously fell ill with a colonic volvulus and was killed. Once the subjects had been trained for 14 sessions, they were allocated to receive either AcbC lesions or sham surgery (0 s: 8 AcbC, 7 sham; 10 s: 8 AcbC, 8 sham; 20 s: 8 AcbC, 8 sham). Sham and AcbC groups were matched for performance preoperatively: within each delay condition, rats were ranked by their rates of responding on the active lever at the end of training, and rats with equivalent levels of performance were randomized to receive sham or AcbC lesion surgery. They were then retested postoperatively on the same task for a further 18 sessions (giving 32 sessions in total), with each rat experiencing the same delay as it had preoperatively. These rats then had their locomotor activity assessed, and were killed and perfused for histology.

2.3.2 Subjects and housing conditions

Subjects were male Lister hooded rats (Harlan-Olac UK Ltd) housed in a temperature-controlled room (minimum 22°C) under a 12:12 h reversed light–dark cycle (lights off 07:30 to 19:30). Subjects were approximately 15 weeks old on arrival at the laboratory and were given a minimum of a week to acclimatize, with free access to food, before experiments began. Experiments took place between 09:00 and 21:00, with individual subjects being tested at a consistent time of day. Subjects had free access to water. During behavioural testing, they were maintained at 85–90% of their free-feeding mass using a restricted feeding regimen. Feeding occurred in the home cages at the end of the experimental day. All procedures were subject to UK Home Office approval (Project Licences PPL 80/1324 and 80/1767) under the Animals (Scientific Procedures) Act 1986.

2.3.3 Excitotoxic lesions of the nucleus accumbens core

Subjects were anaesthetized with Avertin (2% w/v 2,2,2-tribromoethanol, 1% w/v 2-methylbutan-2-ol, and 8% v/v ethanol in phosphate-buffered saline, sterilized by filtration, 10 ml/kg i.p.) and placed in a Kopf or Stoelting stereotaxic frame (David Kopf Instruments, Tujunga, California, USA; Stoelting Co., Wood Dale, Illinois, USA) fitted with atraumatic ear bars (Figure 23). The skull was exposed and a dental drill was used to remove the bone directly above the injection and cannulation sites. The dura mater was broken with the tip of a hypodermic needle, avoiding damage to underlying venous sinuses. Excitotoxic lesions of the AcbC were made by injecting 0.5 µl of 0.09 M quinolinic acid (Sigma, UK) through a glass micropipette at coordinates 1.2 mm anterior to bregma, ±1.8 mm from the midline, and 7.1 mm below the skull surface at bregma; the incisor bar was 3.3 mm below the interaural line (Paxinos & Watson, 1998). The toxin had been dissolved in 0.1 M phosphate buffer (composition 0.07 M Na$_2$HPO$_4$, 0.028 M NaH$_2$PO$_4$ in double-distilled water, sterilized by filtration) and adjusted with NaOH to a final pH of 7.2–7.4. Toxin was injected over 3 min and the micropipette was left in place for 2 min following injections. Sham lesions were made in the same manner except that vehicle was infused. At the end of the operation, animals were given 15 ml/kg of sterile 5% w/v glucose, 0.9% w/v sodium chloride intraperitoneally. They were given a week to recover, with free access to food, and were handled regularly. Any instances of postoperative constipation were

with free access to food, and were handled regularly. Any instances of postoperative constipation were treated with liquid paraffin orally and rectally. At the end of this period, food restriction commenced or was resumed.
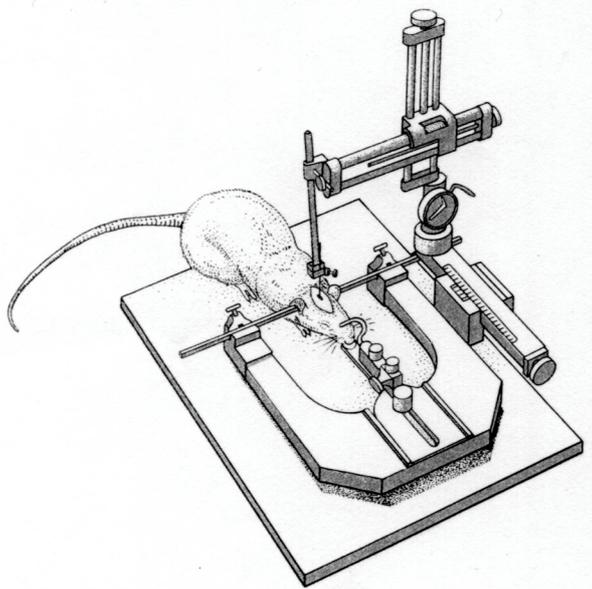
Figure 23: Stereotaxic frame
From Carlson (1991).

2.3.4 Behavioural apparatus

Behavioural testing was conducted in one of two types of operant chamber of identical configuration (from Med Associates Inc, Georgia, Vermont, USA, or Paul Fray Ltd, Cambridge, UK; Figure 24). Each chamber was fitted with a 2.8 W overhead house light and two retractable levers on either side of an alcove fitted with an infrared photodiode to detect head entry. Sucrose pellets (45 mg, Rodent Diet Formula P, Noyes, Lancaster, New Hampshire, USA) could be delivered into the alcove. The chambers were enclosed within sound-attenuating boxes fitted with fans to provide air circulation. The apparatus was controlled by software written by RNC in C++ (Stroustrup, 1986) using the Whisker control system (Cardinal, 2000; Cardinal & Aitken, 2001).

2.3.5 Overview of the Whisker control system

Whisker (Cardinal, 2000; Cardinal & Aitken, 2001) is a research control system written in C++ (Stroustrup, 1986). It uses a client–server architecture, in which a server program controls all external input/output hardware (including, if necessary, sound and video devices) and provides a text-based interface to multiple simultaneous client programs via the transmission control protocol/internet protocol (TCP/IP) system. The server runs on 32-bit Microsoft Windows platforms as a real-time process. It polls hardware at 1 kHz and provides an event-driven environment to the clients, which implement behavioural tasks. The multimedia edition of the server has a memory footprint of approximately 11 Mb and runs to ~45,000 lines of code.

2.3.6 Instrumental conditioning with delayed reinforcement

A variety of free-operant schedules may be used to assess instrumental acquisition with delayed reinforcement (Dickinson *et al.*, 1992). The simplest possible free-operant schedule was used in the present experiments: each response scheduled a reinforcer after the programmed delay (Figure 22, p.53). In such a schedule, if the subject responds during the delay, the experienced response–reinforcer delay will not match the programmed delay (as the second response is temporally close to the first reinforcer). However, this schedule has the advantage that the response–reinforcer contingency is constant (every response does in fact cause the delivery of reinforcement) and the reinforcement rate is not constrained (Dickinson *et al.*, 1992). So that responding could be attributed to the instrumental response–reinforcer contingency, rather than the effects of general activity or reinforcement itself, respond-

a

Figure 24: Operant chamber
(a) Sketch of the operant chambers used in several tasks.
(b) Photograph of such an operant chamber (Med Associates variety). **(c)** Close-up of a rat in the chamber.
From Cardinal (2001). Although one is on in these photographs, stimulus lights were not used in the present experiment.
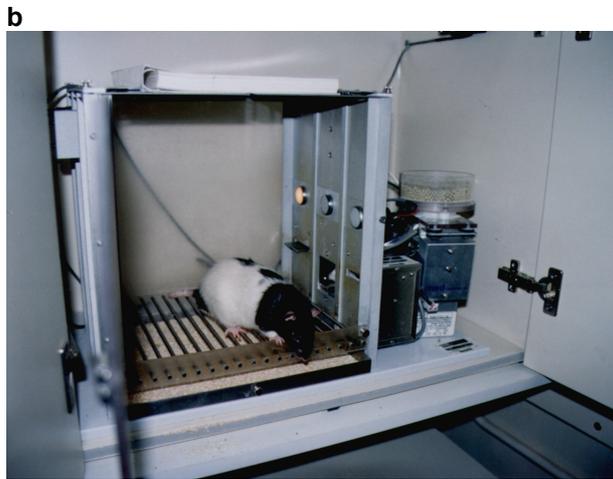
b

c

ing on the active lever was compared to responding on a control lever that had no programmed consequence. Different groups of lesioned and sham-operated subjects were trained using different delays; the delay was consistent for every subject. Delays of 0, 10, and 20 s were used. The schedule was implemented in the SimpleSchedules program (Cardinal, 2002b).

Alternative free-operant schedules for this purpose exist, such as one in which the first response sets up reinforcement, and a subsequent response made before the reinforcer is delivered postpones reinforcement, in order to keep the delay between the last response and the reinforcer constant (known as a tandem fixed-ratio-1 differential-reinforcement-of-other-behaviour or FR-1-DRO schedule). However, the tandem FR-1-DRO schedule constrains the maximum rate of reinforcement, which also decreases as the delay being used increases. Furthermore, it does not hold constant the probability of reinforcement given a response, and it introduces two opposing contingencies: some responses make reinforcement more likely, while others (those during the delay) make it less likely (Dickinson *et al.*, 1992). Therefore, this schedule was not used. Similarly, the acquisition of instrumental responding with delayed reinforcement may be assessed with discrete-trial tasks. For example, two levers could be presented in trials occurring at fixed intervals, the levers could be retracted when a response had been made, and responding on one lever could be reinforced after a delay, taking care to avoid a differential Pavlovian contingency between presentation or retraction of one lever and reinforcement, since responding might then be due to Pavlovian conditioning (autoshaping; Brown & Jenkins, 1968; Williams & Williams, 1969) rather than the instrumental contingency. However, this discrete-trial schedule would also divide up the session explicitly into response–food delays and food–response (intertrial) times, a process that might aid learning and/or be affected by the lesion. Furthermore, there is prior evidence that AcbC lesions impair rats' ability to choose a delayed reward over an immediate reward in the

discrete-trial situation (Cardinal *et al.*, 2001). Therefore, to address the more general question of whether the AcbC is required to acquire instrumental responding with delayed reinforcement, a free-operant schedule was chosen instead; this may mimic best the real-life problem of relating actions to their outcomes with no explicit demarcation of when a response had been made or when a response was permissible.

Immediately after subjects were placed in the operant chamber, the sessions began. The houselight was illuminated, and remained on for each 30-min session. Two levers were extended into the chamber. All lever responses were first "debounced" to 10 ms (i.e. if a response occurred within 10 ms of a previous valid response it was attributed to mechanical bounce and ignored). Other than this, all lever presses and nosepokes into the food alcove were recorded. Responding on the left (active) lever caused a single pellet to be delivered following a delay, under an FR-1 schedule (Figure 22). To attribute acquisition of a lever-press response to the instrumental contingency, it is also necessary to control for the effects of reinforcer delivery itself (Dickinson *et al.*, 1992); therefore, responding on the active lever was compared to responding on the right (inactive) lever, which had no programmed consequence. To minimize any potential contribution of conditioned reinforcement to the task, no explicit signals were associated with pellet delivery other than the noise of the pellet dispenser apparatus.

2.3.7 Locomotor activity in a novel environment

Since general activity levels might influence instrumental responding, locomotor activity was also measured, using wire mesh cages, 25 (W) × 40 (D) × 18 (H) cm, equipped with two horizontal photocell beams situated 1 cm from the floor that enabled movements along the long axis of the cage to be registered. The apparatus was controlled by software written by RNC in Arachnid (Paul Fray Ltd, Cambridge), a real-time extension to BBC BASIC V running on an Acorn Archimedes series computer. Subjects were placed in these cages, which were initially unfamiliar to them, and their activity was recorded for 2 h. All animals were tested in the food-deprived state. Locomotor hyperactivity and reduced weight gain have previously been part of the phenotype of AcbC-lesioned rats, though without alterations in the consumption of the reinforcer used in the present experiments (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001).

2.3.8 Matching of response distribution to reinforcer magnitude distribution on a concurrent schedule

Subjects were trained in 30-min sessions to respond on both levers separately under interval schedules of reinforcement. The two levers were designated A and B; these were counterbalanced left/right (thus, for half the subjects in each group, lever A was the lever reinforced previously in the delay task; for the other half, it was the lever previously unreinforced). As before, responses were debounced to 10 ms. Training and testing proceeded according to Table 4. Random-interval-*x*-second (RI-*x*) schedules were implemented by having a clock tick once a second; each tick set up reinforcement with a probability $p = {}^{1}/_{x}$. Once reinforcement had been set up for a schedule, the next response caused reinforcement to be delivered. Multiple pellets were delivered 0.5 s apart. For concurrent RI schedules, a 2 s changeover delay (COD) was imposed to discourage frequent switching between schedules (Shull & Pliskoff, 1967; Herrnstein, 1970; Williams, 1994; Shahan & Lattal, 1998). The COD was implemented as follows: if a subject pressed lever B, it could only be reinforced if more than 2 s had elapsed since it last pressed lever A (and vice versa). The RI schedules could still set up reinforcement during the COD, but the subject could not earn that reinforcement until the COD had elapsed. The schedule was implemented in the SimpleSchedules program (Cardinal, 2002b).

Day	Condition	ƒA	Lever A	Lever B
1	One-lever training	—	RI-2s, 1-pellet reinforcer	absent
2	One-lever training	—	absent	RI-2s, 1-pellet reinforcer
3	One-lever training	—	RI-15s, 1-pellet reinforcer	absent
4	One-lever training	—	absent	RI-15s, 1-pellet reinforcer
5	One-lever training	—	RI-30s, 1-pellet reinforcer	absent
6	One-lever training	—	absent	RI-30s, 1-pellet reinforcer
7	Two-lever training	0.5	RI-30s, 1-pellet reinforcer	RI-30s, 1-pellet reinforcer
8–11	1:1 magnitude	0.5	RI-60s, 1-pellet reinforcer	RI-60s, 1-pellet reinforcer
12–19	4:1 magnitude	0.8	RI-60s, 4-pellet reinforcer	RI-60s, 1-pellet reinforcer
20–27	1:4 magnitude	0.2	RI-60s, 1-pellet reinforcer	RI-60s, 4-pellet reinforcer

Table 4: Training and testing schedule for reinforcer magnitude matching task

In Experiment 1B, Subjects were trained to respond on two levers (designated A and B) separately and then concurrently under interval schedules of reinforcement. In sessions 8–27, their preference for reinforcers of different magnitudes was assessed. The third column, labelled "ƒA", indicates the fraction of responses that would be allocated to lever A [i.e. A/(A+B)] were the subject to obey the matching law (Herrnstein, 1961). All concurrent (two-lever) schedules were subject to a 2 s changeover delay (COD), described in the Methods.

2.3.9 Histology

Rats were deeply anaesthetized with pentobarbitone sodium (200 mg/ml, minimum of 1.5 ml i.p.) and perfused transcardially with 0.01 M phosphate-buffered saline (PBS) followed by 4% paraformaldehyde in PBS. Their brains were removed and postfixed in paraformaldehyde before being dehydrated in 20% sucrose for cryoprotection. The brains were sectioned coronally at 60 μm thickness on a freezing microtome and every third section mounted on chromium potassium sulphate/gelatin-coated glass microscope slides and allowed to dry. Sections were passed through a series of ethanol solutions of descending concentration (3 minutes in each of 100%, 95%, and 70% v/v ethanol in water) and stained for ~5 min with cresyl violet. The stain comprises 0.05% w/v aqueous cresyl violet (Raymond A. Lamb Ltd, Eastbourne, UK), 2 mM acetic acid, and 5 mM formic acid in water. Following staining, sections were rinsed in water and 70% ethanol before being differentiated in 95% ethanol. Finally, they were dehydrated and delipidated in 100% ethanol and Histoclear (National Diagnostics, UK) before being cover-slipped using DePeX mounting medium (BDH, UK) and allowed to dry. The sections were used to verify cannula and lesion placement and assess the extent of lesion-induced neuronal loss. Lesions were detectable as the absence of visible neurons (cell bodies of the order of 100 μm in diameter with a characteristic shape and appearance), often associated with a degree of tissue collapse (sometimes with consequent ventricular expansion when the lesion was adjacent to a ventricle) and gliosis (visible as the presence of smaller, densely staining cells).

2.3.10 Data analysis

Data collected by the chamber control programs were imported into a relational database (Microsoft Access 97) for case selection and analysed with SPSS 11, according to principles developed in Aitken & Cardinal (2006) and Cardinal & Aitken (2006). Figures were created with SigmaPlot 2001/v7 and Adobe Illustrator 8. All graphs show group means and error bars are ±1 standard error of the mean (SEM) unless otherwise stated. Count data (lever presses and locomotor activity counts), for which variance increases with the mean, were subjected to a square-root transformation prior to any analysis (Howell, 1997). Homogeneity of variance was verified using Levene's test (Levene, 1960). General linear models are described as *dependent variable* $= A_2 \times B_{cov} \times (C_5 \times D_{cov} \times S)$ where A is a between-subjects factor with two levels, B is a between-subjects covariate, C is a within-subjects factor with five levels, and D is a within-subjects covariate; S denotes subjects in designs involving within-subjects factors (Keppel, 1982). For repeated measures analyses, Mauchly's test of sphericity of the covariance matrix was applied (Mauchly,

1940) and the degrees of freedom corrected to more conservative values using the Huynh–Feldt epsilon $\tilde{\varepsilon}$ for any terms involving factors in which the sphericity assumption was violated (Huynh & Feldt, 1970).

2.4 RESULTS

In Experiment 1, rats received excitotoxic lesions of the AcbC or sham lesions, and were then tested on an instrumental free-operant acquisition task with delayed reinforcement (Experiment 1A; see Methods) and subsequently a reinforcer magnitude discrimination task (Experiment 1B). In Experiment 2, naïve rats were trained on the free-operant task for delayed reinforcement; AcbC lesions were then made and the rats were retested.

2.4.1 Histology

In Experiment 1, there were two postoperative deaths. Histological analysis revealed that the lesions were incomplete or encroached significantly on neighbouring structures in four subjects. These subjects were excluded; final group numbers were therefore 8 (sham, 0 s delay) [2], 6 (AcbC, 0 s delay) [3], 8 (sham, 10 s delay) [4], 7 (AcbC, 10 s delay) [5], 8 (sham, 20 s delay) [6], and 7 (AcbC, 20 s delay) [7]. In Experiment 2, one rat spontaneously fell ill with a colonic volvulus during preoperative training and was killed, and there were three postoperative deaths. Lesions were incomplete or too extensive in seven subjects; final group numbers were therefore 7 (sham, 0 s delay) [8], 5 (AcbC, 0 s delay) [9], 8 (sham, 10 s delay) [10], 4 (AcbC, 10 s delay) [11], 8 (sham, 20 s delay) [12], and 5 (AcbC, 20 s delay) [13].

Lesions of the AcbC encompassed most of the core subregion; neuronal loss and associated gliosis extended in an anteroposterior direction from approximately 2.7 mm to 0.5 mm anterior to bregma, and did not extend ventrally or caudally into the ventral pallidum or olfactory tubercle. Damage to the ventromedial caudate–putamen was occasionally seen; damage to AcbSh was restricted to the lateral edge of the dorsal shell. Schematics of the lesions are shown in Figure 25. Photomicrographs of one lesion are shown in Figure 26, and are similar to lesions with identical parameters that have been presented before (Parkinson *et al.*, 1999a; Cardinal, 2001).

---

[2] Experiment 1, sham, 0 s delay: final subjects O1, O2, O3, O4, O5, O6, O7, O8 (*n* = 8).
[3] Experiment 1, AcbC, 0 s delay: final subjects O10, O13, O15, O16, O17, O18 (*n* = 6).
[4] Experiment 1, sham, 10 s delay: final subjects O19, O20, O21, O22, O23, O24, O25, O26 (*n* = 8).
[5] Experiment 1, AcbC, 10 s delay: final subjects O27, O29, O30, O31, O32, O33, O34 (*n* = 7).
[6] Experiment 1, sham, 20 s delay: final subjects O35, O36, O37, O38, O39, O40, O41, O42 (*n* = 8).
[7] Experiment 1, AcbC, 20 s delay: final subjects O44, O45, O46, O47, O48, O49, O50 (*n* = 7).
[8] Experiment 2, sham, 0 s delay: final subjects Q4, Q6, Q8, Q9, Q13, Q15, Q16 (*n* = 7).
[9] Experiment 2, AcbC, 0 s delay: final subjects Q1, Q2, Q7, Q10, Q14 (*n* = 5).
[10] Experiment 2, sham, 10 s delay: final subjects Q18, Q20, Q21, Q23, Q24, Q26, Q31, Q32 (*n* = 8).
[11] Experiment 2, AcbC, 10 s delay: final subjects Q22, Q27, Q28, Q30 (*n* = 4).
[12] Experiment 2, sham, 20 s delay: final subjects Q33, Q34, Q39, Q41, Q43, Q44, Q45, Q46 (*n* = 8).
[13] Experiment 2, AcbC, 20 s delay: final subjects Q35, Q37, Q38, Q40, Q47 (*n* = 5).
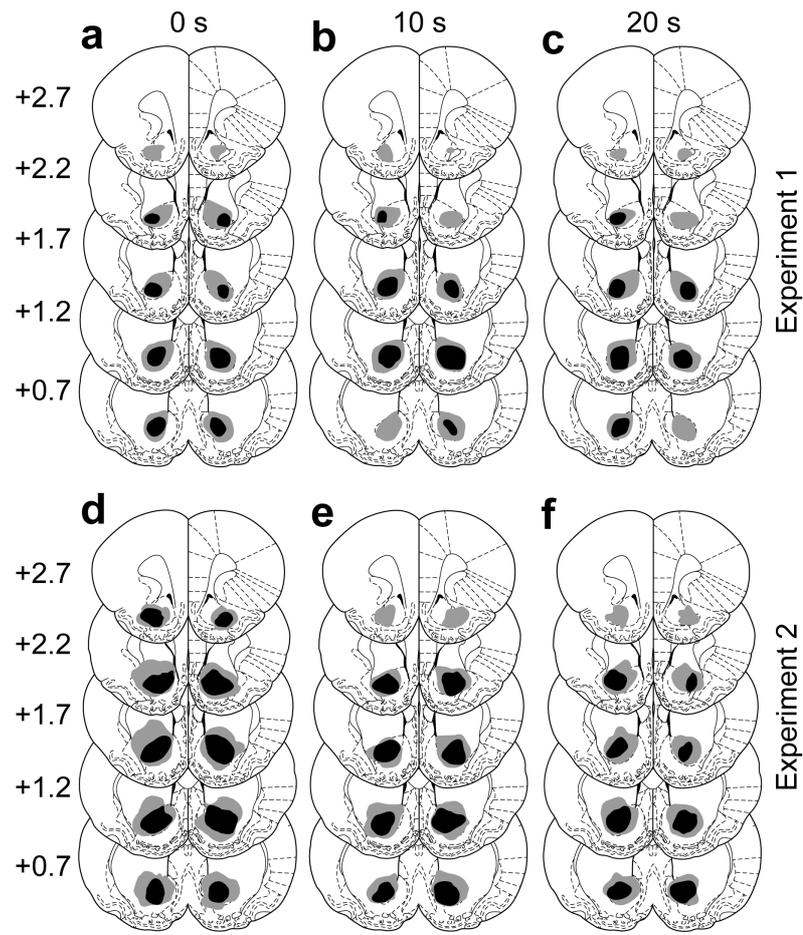
Figure 25: Schematic of lesions of the AcbC

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 (from top to bottom) +2.7, +2.2, +1.7, +1.2, and +0.7 mm relative to bregma. Diagrams are modified from Paxinos & Watson (1998). Panels **a–c** correspond to Experiment 1, in which lesions were made before training; panels **d–f** correspond to Experiment 2, in which lesions were made after initial training. Panels **a & d** show groups trained with no delays; panels **b & e** show groups trained with 10 s delays; panels **c & f** show groups trained with 20 s delays.
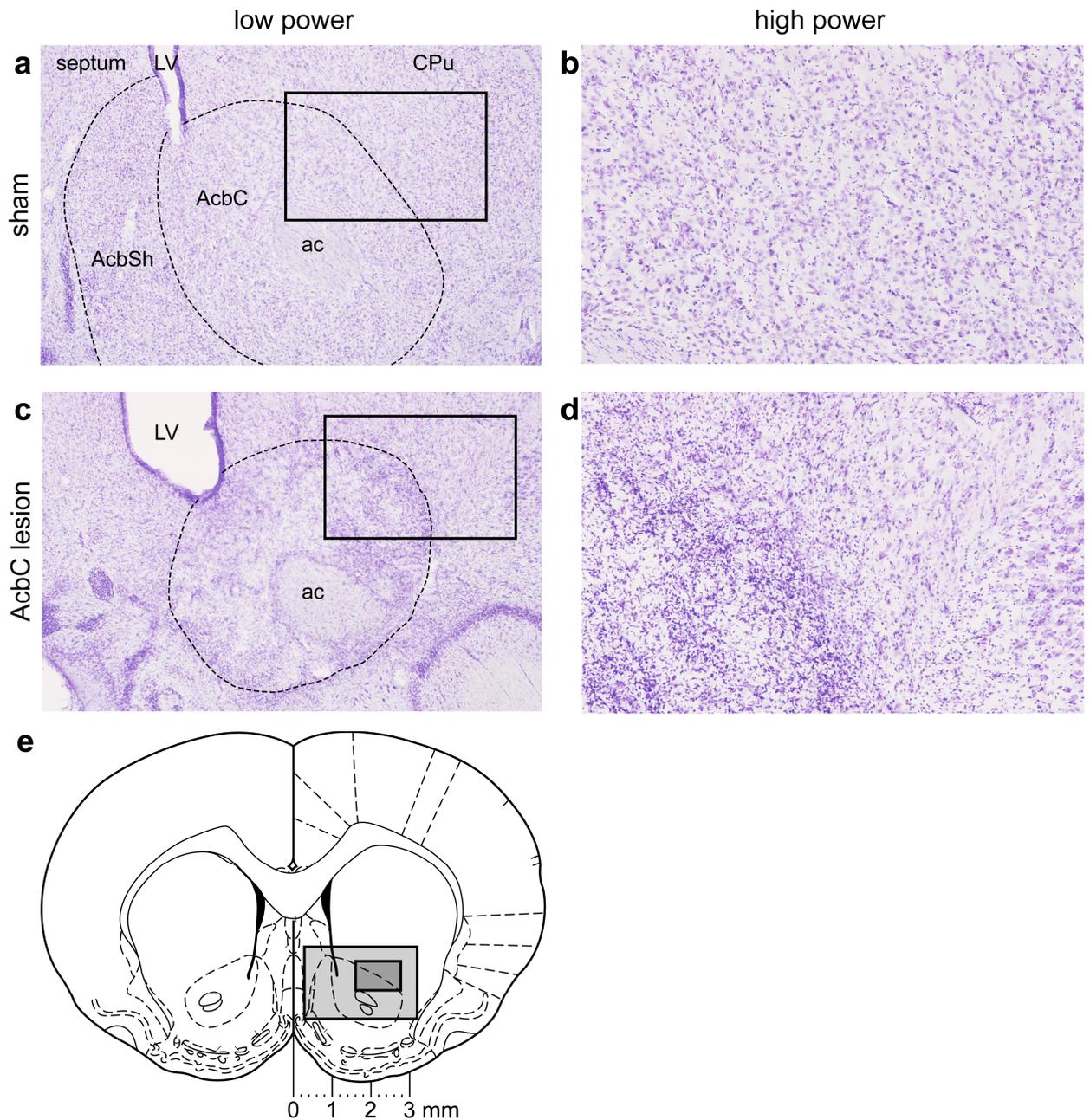
Figure 26: Photomicrographs of lesions of the AcbC

Lesions of the AcbC: photomicrographs of sections ~1.2mm anterior to bregma, stained with cresyl violet. **(a)** Sham-operated rat, low-magnification view, right hemisphere (medial to the left). LV, lateral ventricle; CPu, caudate/putamen; AcbSh, nucleus accumbens shell; AcbC, nucleus accumbens core; ac, anterior commissure. The box marks the area magnified in (b). **(b)** Sham-operated rat, high-magnification view. Cresyl violet is basic and stains for Nissl substance, primarily nucleic acids (DNA and RNA); it therefore stains cytoplasmic rough endoplasmic reticulum, nuclei, and nucleoli. Individual neuronal nuclei are visible (circles ~10 μm in diameter). **(c)** AcbC-lesioned rat, low-magnification view. Dotted lines show the approximate extent of the lesion. There is some tissue collapse within the lesion and the lateral ventricle is slightly expanded. The box marks the area magnified in (d). **(d)** AcbC-lesioned rat, high-magnification view. In the region of the lesion, neurons have been replaced by smaller, densely staining cells, indicating gliosis. **(e)** Coronal diagram of the rat brain at the same anteroposterior level (Paxinos & Watson, 1998), with scale. The light grey box indicates approximately the region shown in (a) and (c); the dark grey box indicates approximately the region shown in (b) and (e).

2.4.2 Acquisition of instrumental responding (Experiment 1A)

The imposition of response–reinforcer delays retarded the acquisition of free-operant lever pressing, in sham-operated rats and in AcbC-lesioned rats (Figure 27). AcbC-lesioned rats responded slightly more than shams on both the active and inactive levers in the absence of response–reinforcers delays, but when such delays were present, AcbC lesions retarded acquisition relative to sham-operated controls (Figure 28).

An overall analysis of variance (ANOVA) using the model $lesion_2 \times delay_3 \times (session_{14} \times lever_2 \times S)$ revealed multiple significant interactions, including lever × delay × lesion ($F_{2,38} = 5.17$, $p = .01$) and session × lever × delay ($F_{6.0,229.1} = 5.47$, $\tilde{\varepsilon} = .464$, $p < .001$), justifying sub-analysis. All six groups learned to respond more on the active lever than the inactive lever ($p \leq .002$, main effect of lever or session × lever interaction for each group alone).

For sham-operated rats, delays reduced the rate of acquisition of the active lever response and reduced the asymptotic level of responding attained (Figure 27a; delay: $F_{2,21} = 11.7$, $p < .001$; $\tilde{\varepsilon} = .276$, $p < .001$; session × delay: $F_{7.2,75.3} = 2.46$, $\tilde{\varepsilon} = .276$, $p = .024$). The presence of a delay also increased responding on the inactive lever slightly (delay: $F_{2,21} = 4.06$, $p = .032$), though not systematically (the 10 s group differed from the 0 s group, $p = .036$, but no other groups differed, $p \geq .153$).

There was a further, delay-dependent impairment in AcbC-lesioned rats, who responded more than shams at 0 s delay but substantially less than shams at 10 s and 20 s delay. As in the case of sham-operated controls, delays reduced the rate of acquisition and the maximum level of responding attained in AcbC-lesioned rats (Figure 27b; delay: $F_{2,17} = 54.6$, $p < .001$; delay × session: $F_{6.9,58.7} = 2.64$, $\tilde{\varepsilon} = .266$, $p = .02$). Responding on the inactive lever was not significantly affected by the delays (maximum $F_{15.8,134.2} = 1.65$, $\tilde{\varepsilon} = .607$, $p = .066$). At 0 s delay, AcbC-lesioned subjects responded more than shams on the active lever (Figure 28a; lesion: $F_{1,12} = 5.30$, $p = .04$) and the inactive lever (lesion: $F_{1,12} = 9.12$, $p = .011$). However, at 10 s delay, AcbC-lesioned rats responded significantly less than shams on the active lever (Figure 28b; lesion: $F_{1,13} = 9.04$, $p = .01$); there was no difference in responding on the inactive lever ($F < 1$, NS). At 20 s delay, again, AcbC-lesioned rats responded significantly less than shams on the active lever (Figure 28c; lesion: $F_{1,13} = 9.87$, $p = .008$) and there was no difference in responding on the inactive lever ($F < 1$, NS).
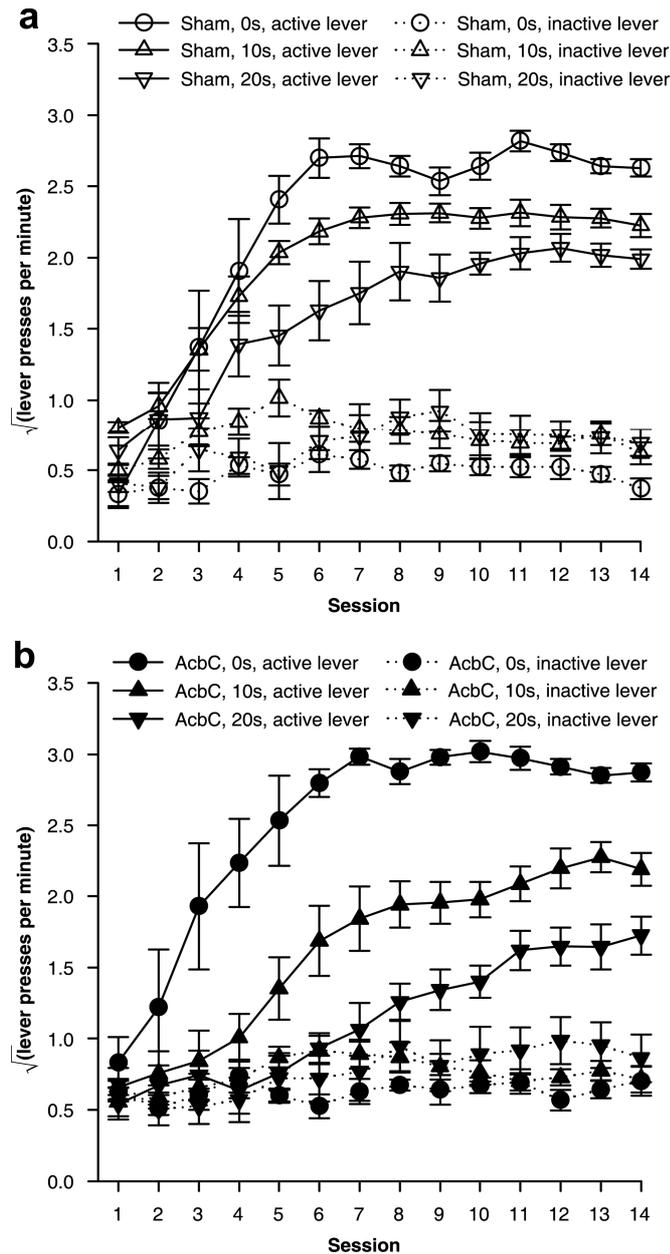
Figure 27: Effects of delays to reinforcement on acquisition of free-operant responding under an FR-1 schedule

Data plotted to show the effects of delays. All groups discriminated between the active and the inactive lever, and delays retarded acquisition of the active lever response in both groups. **(a)** Responding of sham-operated control rats, under all three response–reinforcer delay conditions. **(b)** Responding of AcbC-lesioned rats under all delay conditions. The next figure replots these data to show the effect of the lesion more clearly.
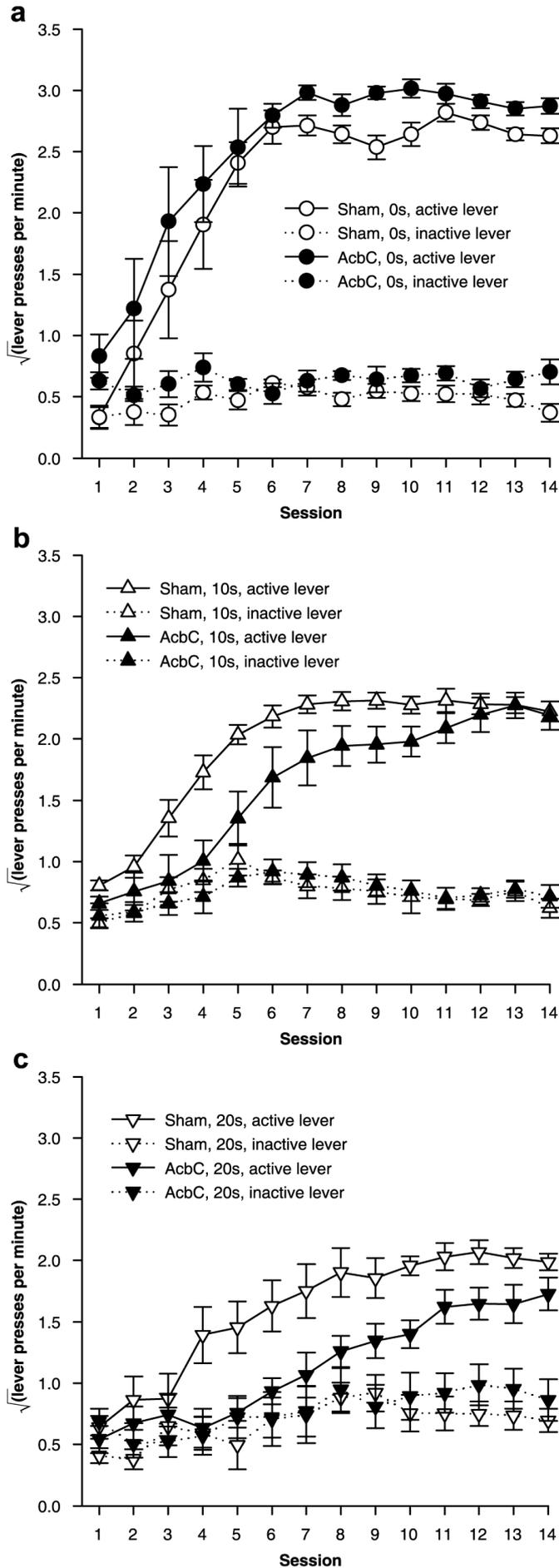
Figure 28: Effect of AcbC lesions on acquisition of free-operant responding with delayed reinforcement

Data plotted to show the effects of AcbC lesions (same data as in the previous figure). There was a delay-dependent impairment in AcbC-lesioned rats, who learned less well than shams only when reinforcement was delayed. **(a)** With a delay of 0 s, AcbC-lesioned rats learned just as well as shams; in fact, they responded more on the active lever than shams did. **(b)** With a 10 s delay, AcbC-lesioned rats were impaired at learning compared to shams. **(c)** With a 20 s delay, the impairment in AcbC-lesioned rats was larger still.

2.4.3 Experienced response–delivery and response–collection delays (Experiment 1A)

For every reinforcer delivered, the active lever response most closely preceding it in time was identified, and the time between that response and delivery of the reinforcer (the "response–delivery delay") was calculated. This time can therefore be equal to or less than the programmed delay, and is only relevant for subjects experiencing non-zero programmed response–reinforcer delays. The response-to-reinforcer-collection ("response–collection") delays were also calculated: for every reinforcer delivered, the response most closely preceding it and the nosepoke most closely following it were identified, and the time between these two events calculated. This time can be shorter or longer than the programmed delay, and is relevant for all subjects.

AcbC-lesioned rats experienced the same response–delivery delays as shams when the programmed delay was 10 s, but experienced longer response–delivery delays when the programmed delay was 20 s (Figure 29a). Similarly, AcbC-lesioned rats experienced the same response–collection delays as shams when the programmed delay was 0 s, slightly but not significantly longer response–collection delays when the programmed delay was 10 s, and significantly longer response–collection delays when the programmed delay was 20 s (Figure 29b). These differences in the *mean* delay experienced by each rat were reflected in differences in the *distribution* of response–delivery and response–collection delays when the programmed delay was non-zero (Figure 29c,d). Since AcbC-lesioned rats experienced slightly longer delays than sham-operated rats, it was necessary to take this into account when establishing the effect of delays on learning, as follows.
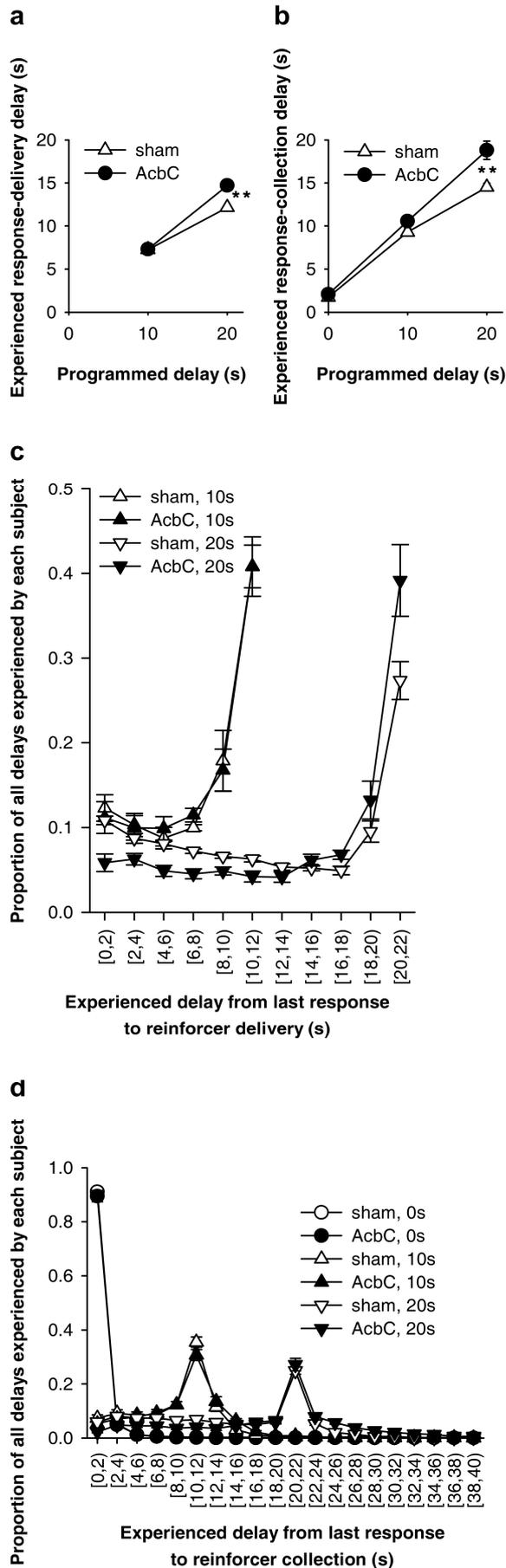
Figure 29: Programmed and experienced delays to reinforcement in AcbC-lesioned and sham-operated rats

AcbC-lesioned rats experienced slightly longer response–delivery delays (the delay between the most recent active lever press and pellet delivery) than shams in the 20 s condition, and slightly longer response–collection delays (the delay between the most recent active lever press and pellet collection) in the 10 s and 20 s conditions. **(a)** Mean experienced response–delivery delays (one value calculated per subject). When the programmed delay was 0 s, reinforcers were delivered immediately so no data are shown. There was a lesion × programmed delay interaction ($F_{1,26}$ = 12.0, p = .002): when the programmed delay was 10 s, the experienced delays did not differ between groups ($F < 1$, NS), but when the programmed delay was 20 s, AcbC-lesioned rats experienced longer response–delivery delays (one-way ANOVA, $F_{1,13}$ = 19.0, ** p = .001). **(b)** Mean experienced response–collection delays (one value calculated per subject). There was a lesion × programmed delay interaction ($F_{2,38}$ = 7.14, p = .002): AcbC-lesioned rats did not experience significantly different delays when the programmed delay was 0 s ($F < 1$, NS) or 10 s ($F_{1,13}$ = 4.52, p = .053), but experienced significantly longer response–collection delays when the programmed delay was 20 s ($F_{1,13}$ = 15.4, ** p = .002). **(c)** Distribution of experienced response–delivery delays. All experienced delays for a given subject were aggregated across all sessions, and the proportion falling into different 2 s ranges were calculated to give one value per range per subject; the graphs show means ± SEMs of these values. The interval notation "[a, b)" indicates that a given delay x falls in the range $a \leq x < b$. There were no differences in the distribution of delays experienced by AcbC-lesioned and sham rats in the 10 s condition (lesion and lesion × range, $Fs < 1$, NS), but in the 20 s condition AcbC-lesioned rats experienced slightly fewer short delays and slightly more long delays (lesion × range, $F_{2.1,27.7}$ = 6.60, $\tilde{\varepsilon}$ = .213, p = .004). **(d)** Distribution of experienced response–collection delays, displayed in the same manner as (c). There were no differences in the distribution of delays experienced by AcbC-lesioned and sham rats in the 0 s condition (lesion and lesion × range, $Fs < 1$, NS). In the 10 s condition, AcbC-lesioned rats experienced a slightly higher proportion of long response–collection delays and a slightly lower proportion of short response–collection delays (lesion, $F_{1,13}$ = 6.36, p = .036, though the lesion × range interaction was not significant, $F_{2.6,34.3}$ = 1.74, $\tilde{\varepsilon}$ = .139, p = .181). Similarly, in the 20 s condition, AcbC-lesioned rats experienced a slightly higher proportion of long response–collection delays and a slightly lower proportion of short response–collection delays than shams (lesion × range, $F_{4.2,54.8}$ = 6.65, $\tilde{\varepsilon}$ = .222, p < .001).

2.4.4 Effect of delays on learning (Experiment 1A)

There was a systematic relationship between the acquisition rate and the programmed delay of reinforcement, and this was altered in AcbC-lesioned rats. Figure 30a replots the rates of responding on the active lever on session 10 of acquisition (cf. Dickinson *et al.*, 1992). Despite the comparatively low power of such an analysis, lever pressing was analysed for this session only using the model lesion$_2$ × delay$_3$. This revealed a significant lesion × delay interaction ($F_{2,38}$ = 12.6, p < .001), which was analysed further. Increasing delays significantly reduced the rate of responding in this session for shams ($F_{2,21}$ = 17.3, p < .001) and AcbC-lesioned rats ($F_{2,17}$ = 54.4, p < .001). AcbC-lesioned rats responded more than shams at zero delay ($F_{1,12}$ = 8.52, p = .013) but less than shams at 10 s delay ($F_{1,13}$ = 4.71, p = .049) and at 20 s delay ($F_{1,13}$ = 17.3, p = .001).

Since the AcbC group experienced slightly longer response–delivery and response–collection delays than shams when the programmed delay was non-zero (Figure 29), it was important to establish whether this effect alone was responsible for the retardation of learning, or whether delays retarded learning in AcbC-lesioned rats over and above any effect to increase the experienced delay. The mean experienced response–collection delay was calculated for each subject, up to and including session 10. The square-root-transformed number of responses on the active lever in session 10 was then analysed using a general linear model of the form lesion$_2$ × experienced delay$_{cov}$. Unlike a standard analysis of covariance (AN-COVA), the factor × covariate interaction term was included in the model. This confirmed that the lesion retarded the acquisition of responding in AcbC-lesioned rats, compared to controls, in a delay-dependent manner, over and above the differences in experienced delay (Figure 30b; lesion × experienced delay: $F_{1,40}$ = 12.4, p = .001).
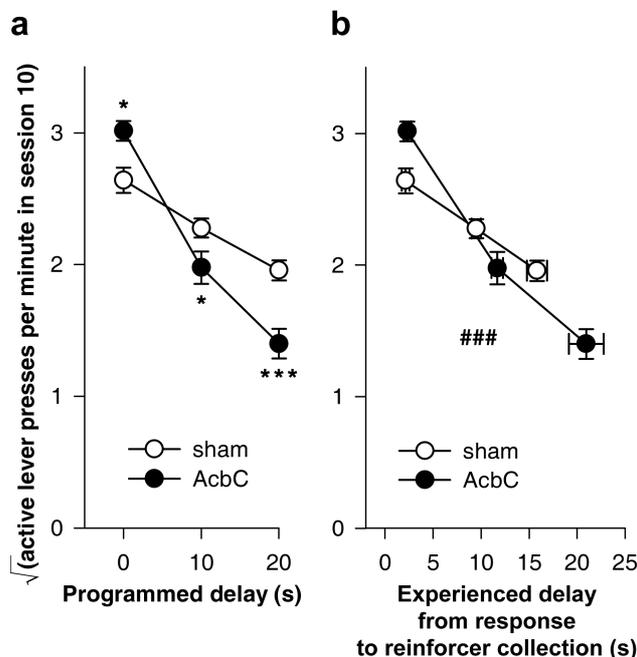
Figure 30: Learning as a function of programmed and experienced delays to reinforcement in AcbC-lesioned and sham-operated rats

The imposition of response–reinforcer delays systematically retarded the acquisition of free-operant instrumental responding, and this relationship was altered in AcbC-lesioned rats, even allowing for differences in experienced response–collection delays. **(a)** The rate of responding on the active lever in session 10 is plotted against the programmed response–reinforcer delay. AcbC-lesioned rats responded more than shams at zero delay (* p = .013), but less than shams at 10 s (* p = .049) and 20 s delay (*** p = .001). **(b)** Responding on the active lever in session 10 plotted against the experienced response-to-reinforcer collection delays for sessions 1–10 (vertical error bars: SEM of the square-root-transformed number of responses in session 10; horizontal error bars: SEM of the experienced response–collection delay, calculated up to and including that session). The gradients of the two lines differed significantly (### p = .001; see text), indicating that the relationship between experienced delays and responding was altered in AcbC-lesioned rats.

2.4.5 Experienced delays and learning on the inactive lever (Experiment 1A)

No such delay-dependent effects were observed for the inactive lever. Experienced inactive-response–delivery delays (calculated across all sessions in the same manner as for the active lever) were much longer and more variable than corresponding delays for the active lever, because subjects responded on

and more variable than corresponding delays for the active lever, because subjects responded on the inactive lever so little. Means ± SEMs were 250 ± 19 s (sham, 0 s), 214 ± 29 s (AcbC, 0 s), 167 ± 23 s (sham, 10 s), 176 ± 33 s (AcbC, 10 s), 229 ± 65 s (sham, 20 s), and 131 ± 37 s (AcbC, 20 s). ANOVA of these data revealed no effects of lesion or programmed delay and no interaction (maximum $F_{1,38} = 1.69$, NS). Experienced inactive-response–collection delays were 252 ± 19 s (sham, 0 s), 217 ± 29 s (AcbC, 0 s), 169 ± 23 s (sham, 10 s), 179 ± 33 s (AcbC, 10 s), 231 ± 65 s (sham, 20 s), and 136 ± 37 s (AcbC, 20 s). Again, ANOVA revealed no effects of lesion or programmed delay and no interaction (maximum $F_{1,38} = 1.61$, NS). When the square-root-transformed number of responses on the inactive lever in session 10 was analysed with the experienced delays up to that point as a predictor, using the model lesion$_2$ × experienced inactive-response–collection delay$_{cov}$ just as for the active lever analysis, there was no lesion × experienced delay interaction ($F < 1$, NS).

2.4.6 Discrimination of relative reinforcer magnitude (Experiment 1B)

Relative preference for two reinforcers may be inferred from the distribution of responses on concurrent VI schedules of reinforcement (Herrnstein, 1961; Herrnstein, 1970; Williams, 1994). According to Herrnstein's matching law (Herrnstein, 1961), if subjects respond on two concurrent schedules A and B delivering reinforcement at rates r_A and r_B respectively, they should allocate their response rates R_A and R_B such that $R_A/(R_A+R_B) = r_A/(r_A+r_B)$. Overmatching is said to occur if subjects prefer the schedule with the higher reinforcement rate more than predicted by the matching law; undermatching is the opposite. Both sham-operated and AcbC-lesioned rats were sensitive to the distribution of reinforcement that they received on two concurrent RI schedules, altering their response allocation accordingly. Subjects preferred the lever on which they received a greater proportion of reinforcement. In general, subjects did not conform to the matching law, but exhibited substantial undermatching; this is common (Williams, 1994). AcbC-lesioned rats exhibited better matching (less undermatching) than shams (Figure 31), suggesting that their sensitivity to the relative magnitudes of the two reinforcers was as good as, or better than, shams'.

To analyse these data, the proportion of pellets delivered by lever A (see Methods, p. 57), and the proportion of responses allocated to lever A, were calculated for each subject for the last session in each of the three programmed reinforcement distribution contingencies (session 11, programmed reinforcement proportion 0.5; session 19, programmed proportion 0.8; session 27, programmed proportion 0.2; see Table 4, p. 58). The analysis used a model of the form response proportion = lesion$_2$ × (experienced reinforcer distribution$_{cov}$ × S); the factor × covariate term was included in the model. Analysis of sham and AcbC groups separately demonstrated that both groups altered their response allocation according to the distribution of reinforcement, i.e. that both groups discriminated the two reinforcers on the basis of their magnitude (effects of reinforcer distribution; sham: $F_{1,47} = 16.6$, $p < .001$; AcbC: $F_{1,39} = 97.2$, $p < .001$). There was also a significant lesion × reinforcer distribution interaction ($F_{1,86} = 5.5$, $p = .021$), indicating that the two groups' matching behaviour differed, with the AcbC-lesioned rats showing better sensitivity to the relative reinforcer magnitude than the shams (Figure 31). These statistical conclusions were not altered by including counterbalancing terms accounting for whether lever A was the left or right lever (the left having been the active lever previously in Experiment 1A), or whether a given rat had been trained with 0, 10, or 20 s delays in Experiment 1A.
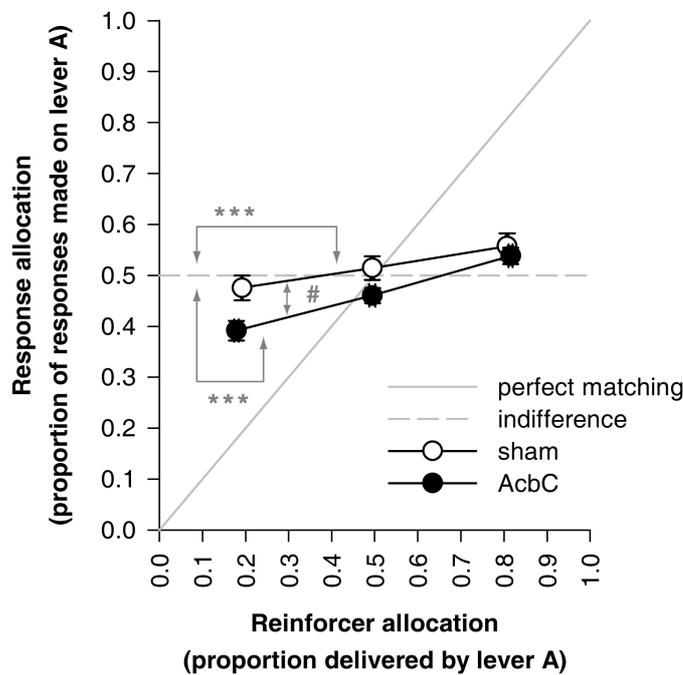
Figure 31: Discrimination of reinforcer magnitude: matching of relative response rate to relative reinforcement rate in AcbC-lesioned and sham-operated rats

AcbC-lesioned rats exhibited better sensitivity to the difference between 1 and 4 food pellets than shams did. Subjects responded on two concurrent RI-60-s schedules, designated A and B, and the reinforcer magnitude for each schedule was varied. Data from the last session of each condition are plotted (sessions 11, 19, and 27; see Table 4, p. 58); programmed reinforcement ratios were 0.2 (1 food pellet on schedule A and 4 pellets on schedule B), 0.5 (1:1 pellets), and 0.8 (4:1 pellets). The abscissa (horizontal axis) shows experienced reinforcement ratios (mean ± SEM); the ordinate (vertical axis) shows response allocation (mean ± SEM). Both groups exhibited substantial undermatching (deviation away from the predictions of the matching law and towards indifference). However, neither group was indifferent to the reinforcement ratio: the sham and AcbC groups both adjusted their response allocation towards the lever delivering the reinforcer with the greater magnitude (*** $p < .001$). Matching was better in AcbC-lesioned rats than in shams (lines of different gradient, # $p = .021$), suggesting that AcbC-lesioned rats were more sensitive to the difference between 1 and 4 food pellets.

2.4.7 Switching behaviour during concurrent schedule performance (Experiment 1B)

Because switching behaviour has the potential to influence behaviour on concurrent schedules (e.g. Shahan & Lattal, 1998), switching probabilities were also analysed. AcbC-lesioned rats were less likely than shams to switch between levers when responding on two identical concurrent RI schedules with a COD of 2 s. Responses on the left and right levers were sequenced for sessions 8–11 (concurrent RI-60s schedules, each delivering a one-pellet reinforcer; see Methods, p. 57, and Table 4, p. 58), and the probabilities of switching from one type of response to another, or repeating the same type of response, were calculated. The switch probabilities were analysed by one-way ANOVA; this revealed an effect of lesion ($F_{1,42}$ = 8.88, $p = .005$). Mean switch probabilities (± SEMs) were 0.41 ± 0.02 (AcbC) and 0.49 ± 0.01 (sham).

2.4.8 Effects of AcbC lesions on performance of a previously learned instrumental response for delayed reinforcement (Experiment 2)

Due to mechanical faults, data from four subjects in session 10 (preoperative) and data from one subject in session 22 (postoperative) were not collected. Both sessions were removed from analysis completely, and data points for those sessions are plotted using the mean and SEM of the remaining unaffected subjects (but not analysed).

Preoperatively, the groups remained matched following later histological selection. Analysis of the last 3 preoperative sessions, using the model lesion intent$_2$ × delay$_3$ × (session$_3$ × lever$_2$ × S), indicated that responding was affected by the delays to reinforcement (delay: $F_{2,31}$ = 5.46, $p = .009$; delay × lever: $F_{2,31}$ = 19.5, $p < .001$), but there were no differences between the groups due to receive AcbC and sham lesions

(terms involving lesion intent: maximum F was for session × lever × lesion intent, $F_{2,62} = 1.844$, NS). As expected, delays reduced the rate of responding on the active lever ($F_{2,31} = 15.6$, $p < .001$) and increased responding on the inactive lever ($F_{2,31} = 8.12$, $p = .001$) preoperatively.

AcbC lesions selectively impaired performance of instrumental responding only when there was a response–reinforcer delay. There was no effect of the lesion on responding under the 0 s delay condition, but in the presence of delays, AcbC lesions impaired performance on the active lever (Figure 32; Figure 33). These conclusions were reached statistically as follows.

Subjects' responding on the relevant lever in the last preoperative session (session 14) was used as a covariate to increase the power of the analysis (Howell, 1997). As expected, there were no significant differences in the covariates themselves between groups due to receive AcbC or sham surgery (terms involving lesion intent for the active lever: Fs < 1, NS; for the inactive lever, lesion intent: $F_{1,31} = 2.99$, $p = .094$; lesion intent × delay: $F < 1$, NS). Analysis of the postoperative sessions, using the model lesion$_2$ × delay$_3$ × (session$_{17}$ × lever$_2$ × session-14-active-lever-responses$_{cov}$ × S), revealed a near-significant lesion × delay × session × lever interaction ($F_{22.4,335.5} = 1.555$, $\tilde{\varepsilon} = .699$, $p = .054$). Furthermore, analysis of postoperative responding on the active lever, using the model lesion$_2$ × delay$_3$ × (session$_{17}$ × session-14-active-lever-responses$_{cov}$ × S), revealed a session × delay × lesion interaction ($F_{17.3,259.5} = 1.98$, $\tilde{\varepsilon} = .541$, $p = .013$) and a delay × lesion interaction ($F_{2,30} = 3.739$, $p = .036$), indicating that the lesion affected responding on the active lever in a delay-dependent manner. In an identical analysis of responding on the inactive lever (using inactive lever responding on session 14 as the covariate), no terms involving lesion were significant (maximum F: lesion, $F_{1,30} = 1.96$, $p = .172$), indicating that the lesion did not affect responding on the inactive lever.

Postoperatively, response–reinforcer delays continued systematically to decrease responding on the active lever, both in shams (Figure 32a; delay: $F_{2,20} = 11.78$, $p < .001$; session × delay: $F_{12.4,124.1} = 2.36$, $\tilde{\varepsilon} = .388$, $p = .008$) and in AcbC-lesioned rats (Figure 32b; delay: $F_{2,11} = 13.9$, $p = .001$). Shams continued to discriminate between the active and inactive lever at all delays (lever: all groups $p \leq .002$; lever × session: all groups $p \leq .003$). AcbC-lesioned rats continued to discriminate at 0 s and 10 s (lever: $p \leq .011$; lever × session: $p \leq .036$), but AcbC-lesioned subjects in the 20 s condition failed to discriminate between the active and inactive levers postoperatively (lever: $F_{1,4} = 1.866$, $p = .244$; lever × session: $F < 1$, NS).

Lesioned subjects responded as much as shams at 0 s delay, but substantially less than shams at 10 s and 20 s delay (Figure 33). Again, analysis was conducted using responding on the relevant lever in session 14 (the last preoperative session) as a covariate. At 0 s, the lesion did not affect responding on the active lever (lesion: $F < 1$, NS; lesion × session: $F_{16,144} = 1.34$, NS). However, at 10 s, AcbC-lesioned rats responded significantly less than shams on the active lever (lesion: $F_{1,9} = 7.08$, $p = .026$; lesion × session: $F_{15.0,135.3} = 3.04$, $\tilde{\varepsilon} = .94$, $p < .001$). Similarly, at 20 s, AcbC-lesioned rats responded less than shams on the active lever (lesion: $F_{1,10} = 6.282$, $p = .031$). There were no differences on responding on the inactive lever at any delay (Fs ≤ 1.31, NS).
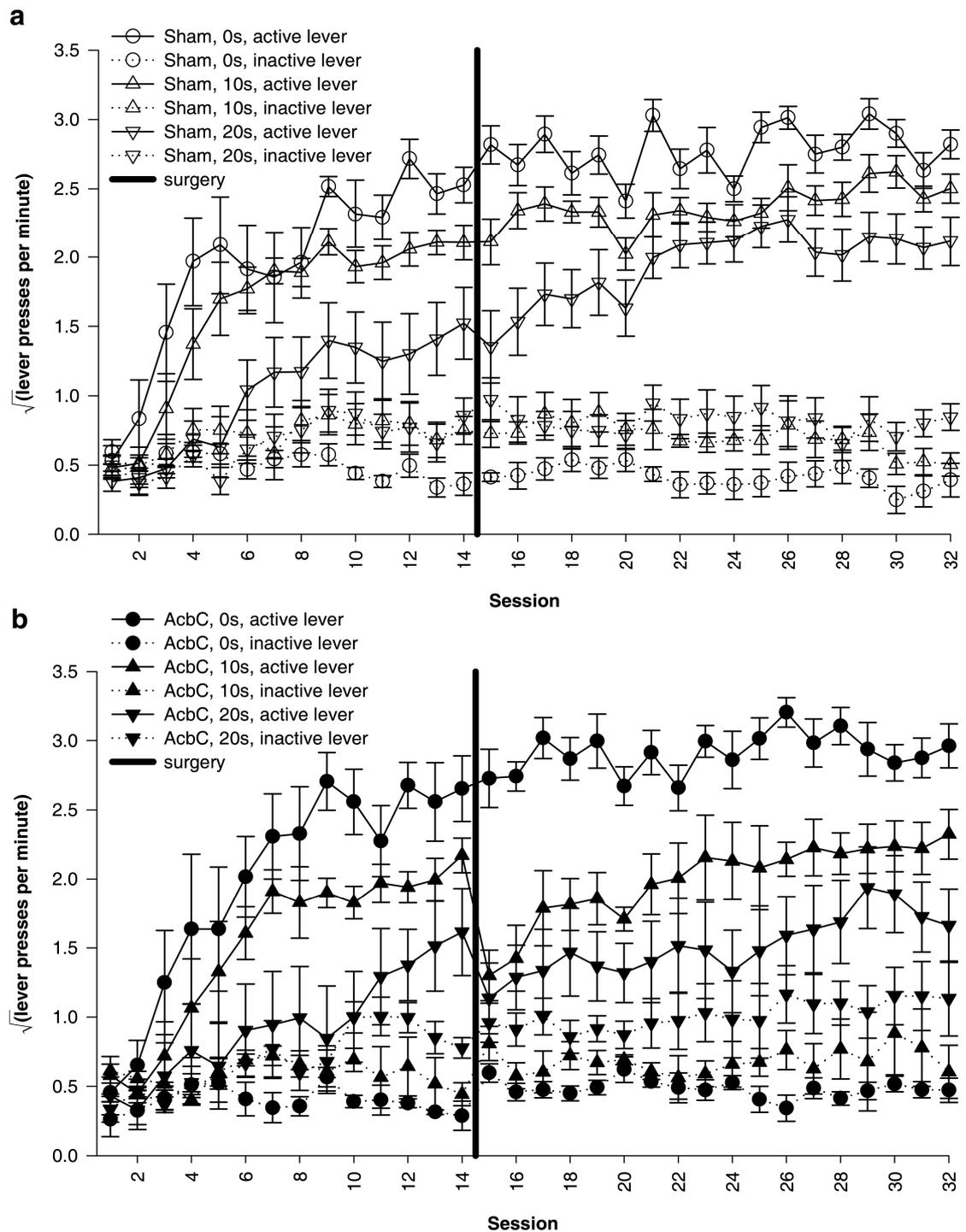
Figure 32: Postoperative performance under an FR-1 schedule for delayed reinforcement

Data plotted to show the effects of delays. All groups discriminated between the active and the inactive lever, and delays retarded acquisition of the active lever response in both groups. Postoperatively, shams' performance was unaltered, as was that of AcbC-lesioned rats in the 0 s delay condition. However, active lever responding was impaired postoperatively in AcbC-lesioned rats in the 10 s and 20 s conditions. **(a)** Responding of sham-operated control rats, under all three response–reinforcer delay conditions. The vertical black line indicates the time of surgery, between testing sessions 14 and 15. **(b)** Responding of AcbC-lesioned rats under all delay conditions. The next figure replots these data to show the effect of the lesion more clearly.
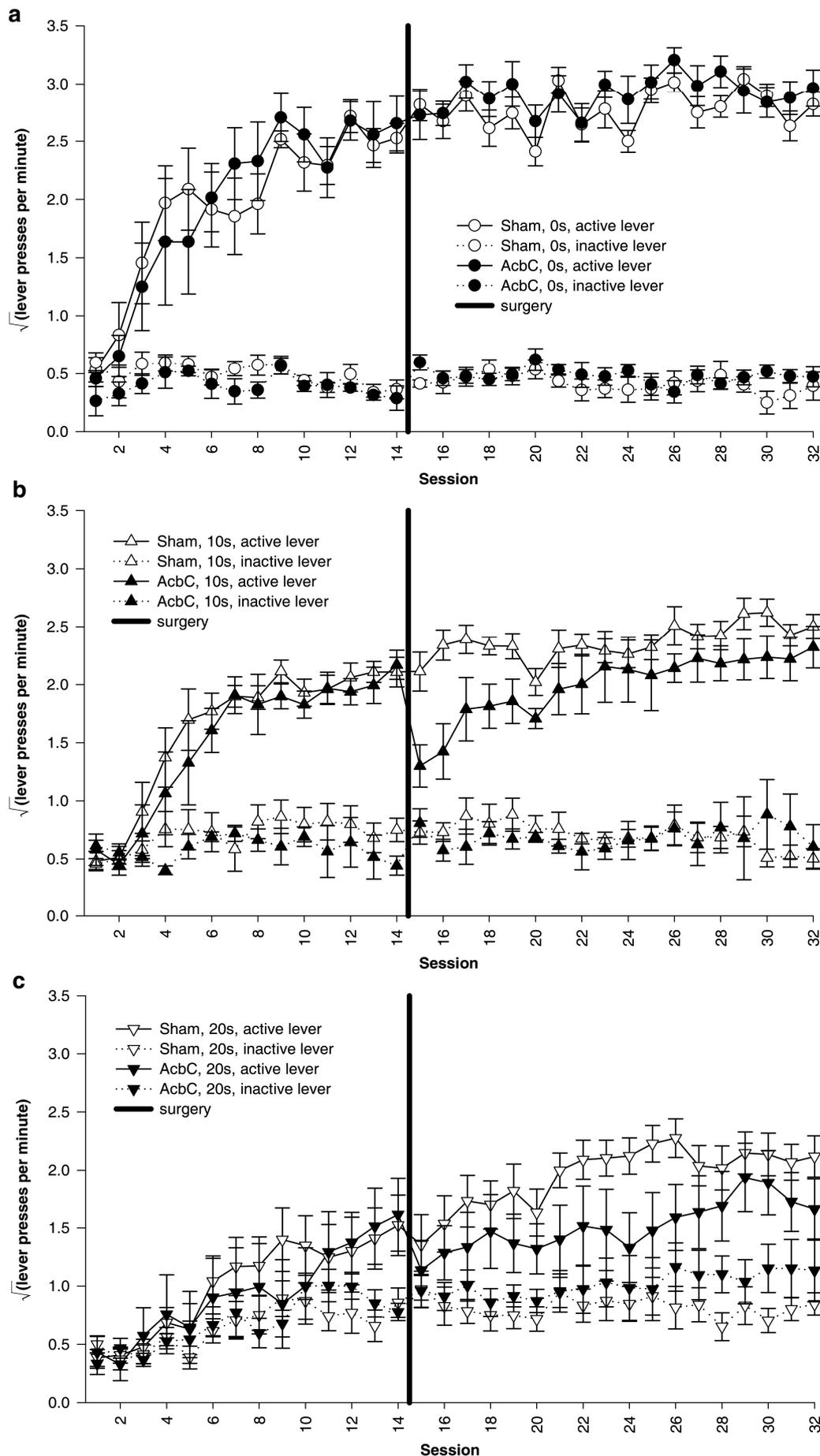
Figure 33: Effect of AcbC lesions on performance of free-operant responding for delayed reinforcement

Data plotted to show the effects of AcbC lesions (same data as in the previous figure). There was a delay-dependent impairment in AcbC-lesioned rats, who were impaired by the lesion only when reinforcement was delayed. **(a)** With a delay of 0 s, AcbC-lesioned rats performed just as well as shams postoperatively. The vertical black line indicates the time of surgery, between testing sessions 14 and 15. **(b)** With a 10 s delay, AcbC-lesioned rats were impaired postoperatively compared to shams. **(c)** With a 20 s delay, the postoperative impairment in AcbC-lesioned rats was larger still, to the extent that their discrimination between active and inactive levers was no longer significant.

2.4.9 Experienced response–delivery and response–collection delays (Experiment 2)

As in Experiment 1, AcbC-lesioned rats experienced the same response–delivery delays as shams when the programmed delay was 10 s, but experienced longer response–delivery delays when the programmed delay was 20 s (Figure 34a). Similarly, AcbC-lesioned rats experienced the same response–collection delays as shams when the programmed delay was 0 s, slightly but not significantly longer response–collection delays when the programmed delay was 10 s, and significantly longer response–collection delays when the programmed delay was 20 s (Figure 34b).
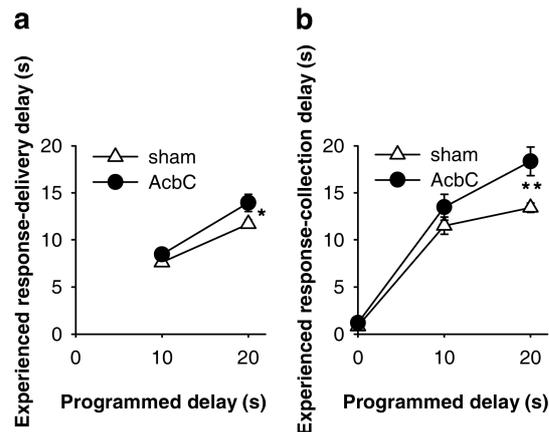
Figure 34: Programmed and experienced delays to reinforcement following AcbC lesions made after initial training

AcbC-lesioned rats experienced slightly longer response–delivery and response–collection delays than shams in the 20 s condition. Lesions were made after initial training; postoperative experienced delays are plotted. (Compare Figure 29, p. 66, in which rats had no preoperative experience of the task.) **(a)** Mean experienced response–delivery delays (one value calculated per subject). When the programmed delay was 0 s, reinforcers were delivered immediately so no data are shown. There were main effects of lesion ($F_{1,21} = 9.14$) and delay ($F_{1,21} = 87.5$, $p < .001$) but no lesion × delay interaction ($F_{1,21} = 1.91$, NS). When the programmed delay was 10 s, the experienced delays did not quite differ significantly between groups ($F_{1,10} = 4.61$, $p = .057$), but when the programmed delay was 20 s, AcbC-lesioned rats experienced longer response–delivery delays ($F_{1,11} = 6.29$, * $p = .029$). **(b)** Mean experienced response–collection delays (one value calculated per subject). There was a lesion × delay interaction ($F_{2,31} = 3.85$, $p = .032$), as well as main effects of lesion ($F_{1,31} = 11.9$, $p = .002$) and delay ($F_{2,31} = 171$, $p < .001$). AcbC-lesioned rats did not experience significantly different delays when the programmed delay was 0 s ($F_{1,10} = 1.74$, NS) or 10 s ($F_{1,10} = 1.49$, NS), but experienced significantly longer response–collection delays when the programmed delay was 20 s ($F_{1,11} = 13.7$, ** $p = .003$).

2.4.10 Relationship between experienced delays and performance (Experiment 2)

There was a systematic relationship between the postoperative response rate and the programmed delay of reinforcement, and this was altered in AcbC-lesioned rats. Figure 35a replots the rates of lever pressing on session 24, the 10th postoperative session (compare Figure 30, p. 67). An analysis using the model lesion$_2$ × programmed delay$_3$ revealed a significant lesion × delay interaction ($F_{2,31} = 5.09$, $p = .012$). In this session, there was no significant effect of delays on shams' performance ($F_{2,20} = 2.15$, $p = .143$), though there was for AcbC-lesioned rats ($F_{2,11} = 9.01$, $p = .005$). There were no significant differences in responding on this session between shams and AcbC-lesioned rats in the 0 s condition ($F_{1,10} = 3.10$, $p = .109$) or the 10 s condition ($F < 1$, NS), but AcbC-lesioned rats responded less at 20 s delay ($F_{1,11} = 6.74$, $p = .025$).

Since the AcbC group experienced slightly longer response–delivery and response–collection delays than shams when the programmed delay was non-zero (Figure 34), as before, the rate of responding in session 24 was analysed as a function of the delays experienced postoperatively. The mean experienced response–collection delay was calculated for postoperative sessions up to and including session 24; the square-root-transformed number of lever presses in session 24 was then analysed using a general linear model of the form lesion$_2$ × experienced delay$_{cov}$, with the factor × covariate interaction term included in the model. This confirmed that the lesion affected responding in AcbC-lesioned rats, compared to controls, in a delay-dependent manner, over and above the postoperative differences in experienced delay (Figure 35b; lesion × experienced delay: $F_{1,33} = 6.53$, $p = .015$).
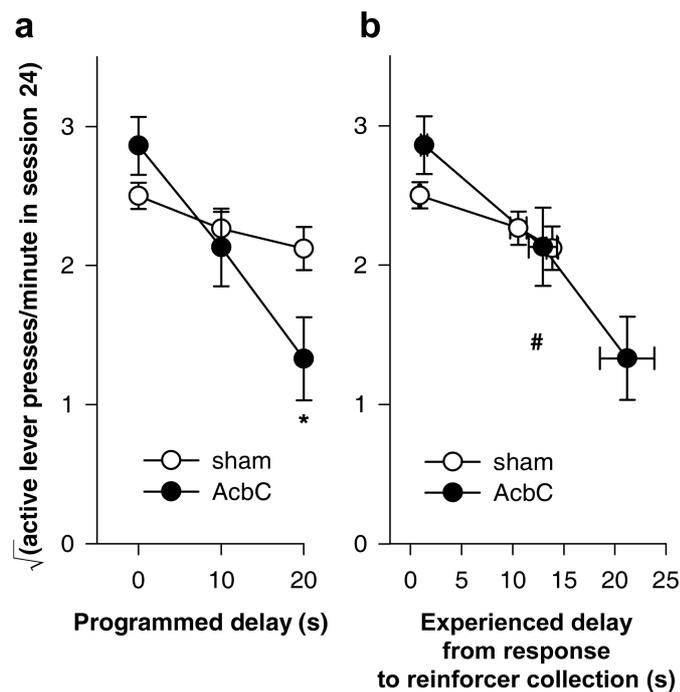
Figure 35: Performance as a function of delays to reinforcement in animals trained preoperatively before sham or AcbC lesions were made

Response–reinforcer delays systematically lowered the rate of free-operant instrumental responding, and this relationship was altered in AcbC-lesioned rats, even allowing for differences in response–collection delays experienced postoperatively. Lesions were made after initial training; postoperative experienced delays and response rates are plotted. (Compare Figure 30, p. 67, in which rats had no preoperative experience of the task.) **(a)** The rate of responding on the active lever in session 24 (the 10th postoperative session; compare Figure 30) is plotted against the programmed response–reinforcer delay. AcbC-lesioned rats responded significantly less than shams in the 20 s delay condition (* $p = .025$). **(b)** Responding on the active lever in session 24 (the 10th postoperative session) plotted against the experienced response-to-reinforcer-collection delays for postoperative sessions up to and including session 24 (vertical error bars: SEM of the square-root-transformed number of responses in session 24; horizontal error bars: SEM of the experienced response–collection delay). The gradients of the two lines differed significantly (# $p = .015$; see text), indicating that the relationship between experienced delays and responding was altered in AcbC-lesioned rats, compared to sham-operated controls.

2.4.11 Locomotor activity and body mass

AcbC-lesioned animals were hyperactive compared to sham-operated controls, and gained less mass then shams across the experiments (Figure 36), consistent with previous results (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001).
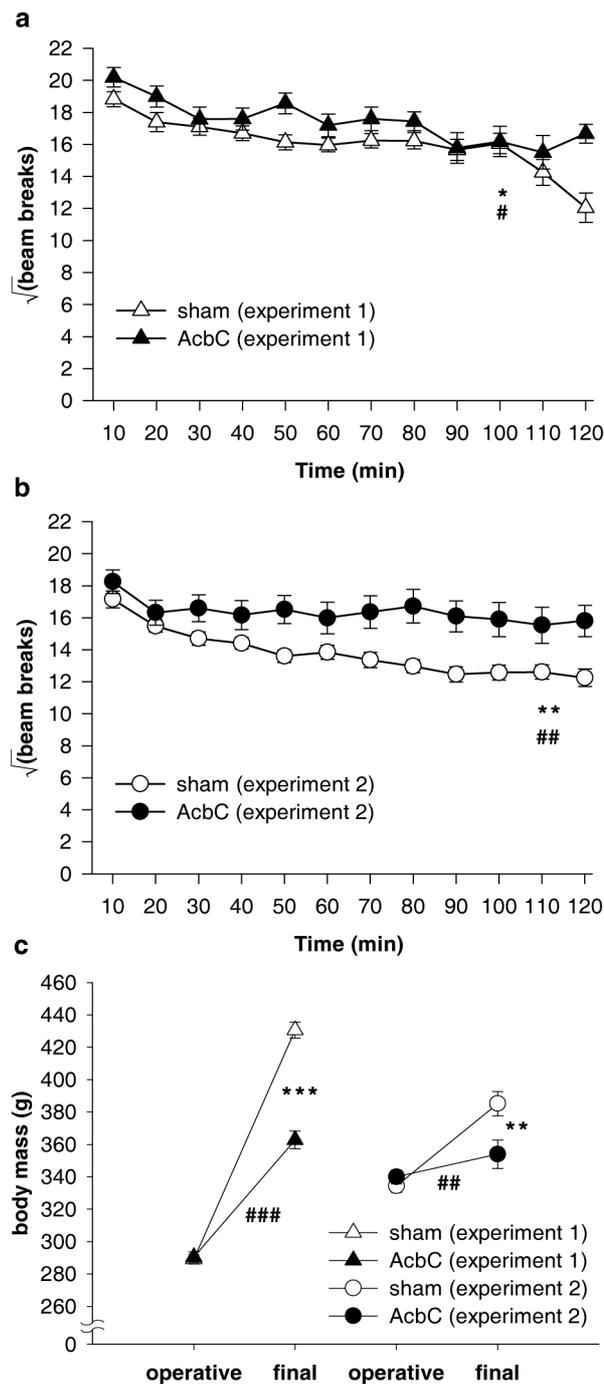
Figure 36: Locomotor activity in a novel environment and body mass in AcbC-lesioned and sham-operated rats

AcbC-lesioned rats were significantly hyperactive compared to sham-operated controls, and gained less weight, in both Experiments 1 & 2. **(a)** Locomotor activity in Experiment 1. Analysis using the model lesion$_2$ × (bin$_{12}$ × S) revealed effects of lesion ($F_{1,42} = 5.12$, * $p = .029$), reflecting hyperactivity in the AcbC group, with additional effects of bin ($F_{5.7,237.9} = 13.3$, $\tilde{\varepsilon} = .515$, $p < .001$), reflecting habituation, and a lesion × bin interaction ($F_{5.7,237.9} = 2.52$, $\tilde{\varepsilon} = .515$, # $p = .024$). **(b)** Locomotor activity in Experiment 2. The same patterns were observed (data from five subjects were not recorded due to a mechanical error; lesion: $F_{1,37} = 9.155$, ** $p = .004$; bin: $F_{9.3,345.2} = 13.5$, $\tilde{\varepsilon} = .848$, $p < .001$; lesion × bin: $F_{9.3,345.2} = 3.18$, $\tilde{\varepsilon} = .848$, ## $p = .001$). **(c)** Preoperative and final body mass in both experiments. Preoperatively, masses did not differ between groups (Experiment 1: $F < 1$, NS; Experiment 2: $F_{1,42} = 1.008$, NS), but in both cases, AcbC-lesioned subjects gained less mass than controls (Experiment 1: lesion × time: $F_{1,41} = 95.9$, ### $p < .001$; group difference at second time point: $F_{1,42} = 88.4$, *** $p < .001$; Experiment 2: lesion × time: $F_{1,42} = 13.53$, ## $p = .001$; group difference at second time point: $F_{1,42} = 7.37$, ** $p = .01$).

2.5 DISCUSSION

These results establish that the AcbC contributes to learning of actions when the outcome is delayed. Lesions of the AcbC did not impair instrumental learning when the reinforcer was delivered immediately, but substantially impaired learning with delayed reinforcement, indicating that the AcbC "bridges" action–outcome delays during learning. Lesions made after learning also impaired performance of the instrumental response in a delay-dependent fashion, indicating that the AcbC also contributes to the performance of actions for delayed reinforcement. Finally, the lesions did not impair the perception of rela-

tive reward magnitude as assessed by responding on identical concurrent interval schedules for reinforcers of different magnitude, suggesting that the impulsive choice previously exhibited by AcbC-lesioned rats (Cardinal *et al.*, 2001) is attributable to deficits in dealing with delays to reinforcement.

2.5.1 Effect of delays on instrumental learning in normal animals

As discussed in Chapter 1, delays have long been known to retard instrumental learning (Grice, 1948; Dickinson *et al.*, 1992). Despite this, normal rats have been shown to acquire free-operant responding with programmed response–reinforcer delays of up to 32 s, or even 64 s if the subjects are pre-exposed to the learning environment (Dickinson *et al.*, 1992). Delays do reduce the asymptotic level of responding (Dickinson *et al.*, 1992), though the reason for this phenomenon is not clear. It may be that when subjects learn a response with a substantial response–reinforcer delay, they never succeed in representing the instrumental action–outcome contingency fully. Alternatively, they may value the delayed reinforcer slightly less; finally, the delay may also retard the acquisition of a procedural stimulus–response habit and this might account for the decrease in asymptotic responding. It is not presently known to what degree responses acquired with a response–reinforcer delay are governed by declarative processes (the action–outcome contingency plus a representation of the instrumental incentive value of the outcome) or procedural mechanisms (stimulus–response habits), both of which are known to influence instrumental responding (Dickinson, 1994; Dickinson & Balleine, 1994); it is similarly not known whether the balance of these two controlling mechanisms differs from that governing responses learned without such a delay.

2.5.2 Effect of AcbC lesions on instrumental learning and performance with or without delays

In the absence of response–reinforcer delays, AcbC-lesioned rats acquired an instrumental response normally, responding even more than sham-operated controls. In contrast, blockade of NMDA glutamate receptors in the AcbC has been shown to retard instrumental learning for food under a VR-2 schedule [in which $P(reinforcer \mid response) \approx 0.5$] (Kelley *et al.*, 1997), as has inhibition or over-stimulation of PKA within the Acb (Baldwin *et al.*, 2002a). Concurrent blockade of NMDA and DA D$_1$ receptors in the AcbC synergistically prevents learning of a VR-2 schedule (Smith-Roe & Kelley, 2000). Once the response has been learned, subsequent performance on this schedule is not impaired by NMDA receptor blockade within the AcbC (Kelley *et al.*, 1997). Furthermore, infusion of a PKA inhibitor (Baldwin *et al.*, 2002a) or a protein synthesis inhibitor (Hernandez *et al.*, 2002) into the AcbC *after* instrumental training sessions impairs subsequent performance, implying that PKA activity and protein synthesis in the AcbC contribute to the consolidation of instrumental behaviour. Thus, manipulation of Acb neurotransmission can affect instrumental learning. However, it is also clear that excitotoxic destruction of the AcbC or even the entire Acb does not impair simple instrumental conditioning to any substantial degree. Rats with Acb or AcbC lesions acquire lever-press responses on sequences of RR schedules [in which $P(reinforcer \mid response)$ typically declines from around 1 to 0.05 over training] at near-normal levels (Corbit *et al.*, 2001; de Borchgrave *et al.*, 2002). In such ratio schedules, where several responses are required to obtain reinforcement, there is no delay between the final response and reinforcement, but there are delays between earlier responses and eventual reinforcement. It is therefore of interest that when differences between AcbC-lesioned rats and shams have been observed, AcbC-lesioned animals have been found to respond somewhat less than shams on such schedules late in training, when the ratio requirement is high (Corbit *et al.*, 2001; de Borchgrave *et al.*, 2002), consistent with the present results. However, lesioned rats are fully sensitive to changes in the instrumental contingency (Balleine & Killcross, 1994; Corbit *et al.*, 2001; de

Borchgrave *et al.*, 2002). The present results indicate that when AcbC-lesioned rats are exposed to a FR-1 schedule for food [$P(reinforcer \mid response) = 1$] in the absence of response–reinforcer delays, they acquire the response at normal rates.

In contrast, when a delay was imposed between responding and reinforcement, AcbC-lesioned rats were impaired relative to sham-operated controls, in a systematic and delay-dependent fashion. The observation that learning was not affected at zero delay rules out a number of explanations of this effect. For example, it cannot be that AcbC-lesioned rats are in some way less motivated for the food *per se*, since they responded normally (in fact, more than shams) when the food was not delayed. Thus although the Acb and its dopaminergic innervation are clearly very important in motivating behaviour (e.g. Ikemoto & Panksepp, 1999; Cardinal *et al.*, 2002a; Salamone & Correa, 2002; Salamone *et al.*, 2003), this is not on its own a sufficient explanation for the present results. An explanation in terms of a rate-dependent impairment is also not tenable, since the AcbC-lesioned rats were capable (in the zero-delay condition) of responding at a level greater than they exhibited in the non-zero-delay conditions. Depletion of Acb DA also impairs rats' ability to work on high-effort schedules, where many, or very forceful, responses are required to obtain a given amount of food (Salamone & Correa, 2002; Salamone *et al.*, 2003). However, in the present experiments the ratio requirement (one response per reinforcer) and the force required per press were both held constant across delays, so this effect cannot explain the present results. Similarly, although AcbC lesions are known to impair the control over behaviour by Pavlovian CSs (Everitt *et al.*, 1991; Parkinson *et al.*, 1999a; Parkinson *et al.*, 2000c; Hall *et al.*, 2001; Cardinal *et al.*, 2002a; 2002b), there was no Pavlovian stimulus that was differentially associated with delayed as opposed to immediate reinforcement in this task, so this cannot explain the present results.

The present results also indicated that when there were programmed delays to reinforcement, AcbC-lesioned animals experienced longer response–reinforcer collection delays, partly due to their failure to collect the reinforcer as promptly as shams. These additional experienced delays probably retarded learning. However, in addition to this effect, there was a further deficit exhibited by AcbC-lesioned rats: even allowing for the longer response–collection delays that they experienced, their instrumental learning was impaired more by delays than that of sham-operated controls. Deficits in learning with delayed reinforcement may account for some of the variability in the effect of AcbC lesions or local pharmacological manipulations on instrumental learning across different schedules.

The fact that pre-exposure to the context improves instrumental learning in normal rats (Dickinson *et al.*, 1992) suggests one possible mechanism by which AcbC lesions might retard learning when delays are present. When a reinforcer arrives, it may be associated either with a preceding response, or with the context. Therefore, in normal animals, pre-exposure to the context may retard the formation of context–reinforcer associations by latent inhibition, or it might serve to retard the formation of associations between irrelevant behaviours and reinforcement. Similarly, non-reinforced exposure to the context forces the subjects to experience a zero-response, zero-reinforcer situation, i.e. $P(outcome \mid no\ action) = 0$. When they are then exposed to the instrumental contingency, such that $P(outcome \mid action) > 0$, this prior experience may enhance their ability to detect the instrumental contingency $\Delta P = P(outcome \mid action) - P(outcome \mid no\ action)$. In one aversive Pavlovian conditioning procedure in which a CS was paired with electric shock, AcbC lesions have been shown to impair conditioning to discrete CSs, but simultaneously to enhance conditioning to contextual (background) CSs (Parkinson *et al.*, 1999b), though not all behavioural paradigms show this effect (Levita *et al.*, 2002; Jongen-Relo *et al.*, 2003). It is therefore possible that enhanced formation of context–reinforcer associations may explain the retardation of response–reinforcer learning in AcbC-lesioned rats in the presence of delays.

The instrumental task used requires animals either to associate their response with the delayed food outcome (an action–outcome association that can be used for goal-directed behaviour), or to strengthen a stimulus–response association (habit) when the reinforcer eventually arrives (Dickinson, 1994; Dickinson & Balleine, 1994). Both mechanisms require the animal to maintain a representation of their past action so it can be reinforced (as a habit) or associated with food when the food finally arrives. This mnemonic requirement is not obviated even if the animal learns to predict the arrival of food using discriminative stimuli, and uses these stimuli to reinforce its responding (conditioned reinforcement): in either case, since the action precedes reinforcement, some trace of past actions or stimuli must persist to be affected by the eventual delivery of food.

A delay-dependent impairment was also seen when AcbC lesions were made after training. This indicates that the AcbC does not only contribute to the learning of a response when there is an action–outcome delay: it also contributes to the performance of a previously learned response. Again, AcbC-lesioned rats were only impaired when that previously learned response was for delayed (and not immediate) reinforcement. Of course, learning of an instrumental response depends upon the animal being able to perform that response; preventing an animal from pressing a lever (a performance deficit) would clearly impair its ability to learn an instrumental response on that lever to obtain food. In the present set of experiments, it is clear that AcbC-lesioned rats were just as able to perform the response itself (to press the active lever and to discriminate it physically from the inactive lever) as controls, as shown by their normal performance in the zero-delay condition, so it is not clear whether the delay-dependent impairments in learning and performance can be attributed to the same process. Again, since responding was unaffected in the zero-delay condition, many alternative interpretations (such as a lack of motivation to work for the food) are ruled out. It may be that AcbC-lesioned rats are impaired at representing a declarative instrumental action–outcome contingency when the outcome is delayed, or in forming or executing a procedural stimulus–response habit when the reinforcing event does not follow the response immediately. It may also be that they represent the action–outcome contingency normally but value the food less because it is delayed, and that this affects responding in a free-operant situation even though there is no alternative reinforcer available.

2.5.3 Discrimination of reinforcer magnitude in AcbC-lesioned rats

Excitotoxic lesions of the whole Acb do not prevent rats from detecting changes in reward value (induced either by altering the concentration of a sucrose reward or by changing the deprivational state of the subject) (Balleine & Killcross, 1994). Such lesions also do not impair rats' ability to respond faster when environmental cues predict the availability of larger rewards (Brown & Bowman, 1995), and nor does inactivation of the Acb with local anaesthetic or blockade of AMPA glutamate receptors in the Acb (Giertler *et al.*, 2004); the effects of intra-Acb NMDA receptor antagonists have varied (Hauber *et al.*, 2000; Giertler *et al.*, 2003; 2005). AcbC-lesioned rats can still discriminate large from small rewards (Cardinal *et al.*, 2003b; 2004). Similarly, DA depletion of the Acb does not affect the ability to discriminate large from small reinforcers (Salamone *et al.*, 1994; Cousins *et al.*, 1996; Salamone *et al.*, 2001), and systemic DA antagonists do not affect the perceived quantity of food as assessed in a psychophysical procedure (Martin-Iverson *et al.*, 1987). The present study extends these findings by demonstrating that excitotoxic AcbC lesions do not impair rats' ability to allocate their responses across two schedules in proportion to the experienced reinforcement rate, even when the two schedules are identical except in the magnitude of the reinforcements they provide, thus demonstrating their sensitivity to reinforcer magnitude is quantitatively no worse than shams'. In this experiment, there was substantial undermatching, but this is common

(Davison & McCarthy, 1988; Williams, 1994) (see also Leon & Gallistel, 1998; Weatherly *et al.*, 2004); differential cues signalling the two rewards might have improved matching but were not used in the present experiments since it is known that AcbC lesions can themselves affect rats' sensitivity to cues signalling reinforcement (Everitt *et al.*, 1991; Parkinson *et al.*, 1999a; Parkinson *et al.*, 2000c; Hall *et al.*, 2001; Cardinal *et al.*, 2002a; 2002b). Given that AcbC-lesioned subjects showed a reduced probability of switching between two identical RI schedules, it may be the case that an enhanced sensitivity to the COD accounts for the better matching exhibited by the AcbC-lesioned rats (Shahan & Lattal, 1998). Alternatively, the lesion may have enhanced reinforcer magnitude discrimination or improved the process by which behaviour allocation is matched to environmental contingencies. In summary, the present results suggest that AcbC damage leads to pathological impulsive choice (preferring a small, immediate reinforcer to a large, delayed reinforcer; Cardinal *et al.*, 2001) not through any relative lack of value of large reinforcers, but through a specific deficit in responding for delayed reinforcement.

2.5.4 Contribution of the AcbC to reinforcement learning

The term "reinforcement learning" simply means learning to act on the basis of reinforcement received; it is a term used in artificial intelligence research (Minsky, 1961) that does not specify the mechanism of such learning (Russell & Norvig, 1995; Haykin, 1999). The present results indicate that the AcbC is a reinforcement learning structure that is critical for instrumental conditioning when outcomes are delayed, consistent with electrophysiological and functional neuroimaging evidence indicating that the ventral striatum responds to recent past actions (Schultz *et al.*, 2000; Cromwell & Schultz, 2003) and to predicted future rewards (Schultz *et al.*, 1992; Miyazaki *et al.*, 1998; Martin & Ono, 2000; Schultz *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; Cromwell & Schultz, 2003; Bjork *et al.*, 2004), and with computational models suggesting a role for the striatum in predicting future primary reinforcement (Houk *et al.*, 1995; Wickens & Kötter, 1995). However, when reward is certain and delivered immediately, the AcbC is not necessary for the acquisition of instrumental responding. The delay-dependent role of the AcbC indicates that it plays a role in allowing actions to be reinforced by bridging action–outcome delays through a representation of past acts or future rewards. Acb lesions have also produced delay-dependent impairments in a delayed-matching-to-position task (Dunnett, 1990; Reading & Dunnett, 1991); their effects on the delayed-matching-to-sample paradigm have also been studied, but a more profound and delay-independent deficit was observed, likely due to differences in the specific task used (Burk & Mair, 2001). Finally, the AcbC is not alone in containing neurons that respond to past actions and future rewards. The dorsal striatum is another such structure (Schultz *et al.*, 2000; Takikawa *et al.*, 2002; Cromwell & Schultz, 2003; Kawagoe *et al.*, 2004); expression of stimulus–response habits requires the dorsal striatum (Packard & McGaugh, 1996; Yin *et al.*, 2004), and the rate at which rats learn an arbitrary response that delivers electrical stimulation to the substantia nigra is correlated with the degree of potentiation of synapses made by cortical afferents onto striatal neurons, a potentiation that requires DA receptors (Reynolds *et al.*, 2001; Reynolds & Wickens, 2002). The prelimbic area of rat PFC is important for the detection of instrumental contingencies and contributes to goal-directed, rather than habitual, action (Balleine & Dickinson, 1998; Corbit & Balleine, 2003). Similarly, the OFC and BLA encode reinforcement information and project to the AcbC, and lesions of these structures can produce impulsive choice (see Kheramin *et al.*, 2002; Mobini *et al.*, 2002; Cardinal *et al.*, 2004; Winstanley *et al.*, 2004b). It is not yet known whether lesions of these structures also impair learning with delayed reinforcement.

2.6 CONCLUSIONS

These experiments have demonstrated that excitotoxic lesions of the AcbC do not prevent rats from learning a simple instrumental response when the reinforcing outcome follows their action immediately. However, AcbC lesions impair rats' ability to learn the same instrumental response when the outcome is delayed. The lesions also impair performance of an instrumental response that was learned preoperatively, but again only when response–reinforcer delays are present. These results suggest that the AcbC makes a specific contribution to reinforcement learning and instrumental performance when reinforcing outcomes do not arrive immediately but are delayed. AcbC dysfunction, which is known to promote impulsive choice, appears to cause rats to be temporally short-sighted, learning preferentially about the proximal consequences of their actions and preferring immediate over delayed rewards.

Chapter 3: Hippocampal lesions facilitate instrumental learning with delayed reinforcement but induce impulsive choice in rats

3.1 ABSTRACT

Background: Animals must frequently act to influence the world even when the reinforcing outcomes of their actions are delayed. Learning with action–outcome delays is a complex problem, and little is known of the neural mechanisms that bridge such delays. When outcomes are delayed, they may be attributed to (or associated with) the action that caused them, or mistakenly attributed to other stimuli, such as the environmental context. Consequently, animals that are poor at forming context–outcome associations might learn action–outcome associations better with delayed reinforcement than normal animals. The hippocampus contributes to the representation of environmental context, being required for aspects of contextual conditioning. It was therefore hypothesized that animals with hippocampal lesions would be better than normal animals at learning to act on the basis of delayed reinforcement. These experiments tested the ability of H-lesioned rats to learn a free-operant instrumental response using delayed reinforcement, and what is potentially a related ability—the ability to exhibit self-controlled choice, or to sacrifice an immediate, small reward in order to obtain a delayed but larger reward.

Results: Rats with sham or excitotoxic hippocampal lesions acquired an instrumental response with different delays (0, 10, or 20 s) between the response and reinforcer delivery. These delays retarded learning in normal rats. H-lesioned rats responded slightly less than sham-operated controls in the absence of delays, but they became better at learning (relative to shams) as the delays increased; delays impaired learning less in H-lesioned rats than in shams. In contrast, lesioned rats exhibited impulsive choice, preferring an immediate, small reward to a delayed, larger reward, even though they preferred the large reward when it was not delayed.

Conclusions: These results support the view that the hippocampus hinders action–outcome learning with delayed outcomes, perhaps because it promotes the formation of context–outcome associations instead. However, although lesioned rats were better at learning with delayed reinforcement, they were worse at choosing it, suggesting that self-controlled choice and learning with delayed reinforcement tax different psychological processes.

3.2 BACKGROUND

When one event or stimulus in the world reliably precedes and predicts another, animals readily learn the predictive relationship, exemplified by Pavlovian conditioning (Pavlov, 1927). Similarly, when an animal's own actions cause (and thus predict) some outcome, animals learn this relationship (an aspect of instrumental or operant conditioning). Frequently, however, antecedent and consequent events are separated in time. When animals act to obtain reinforcement, the final outcomes do not always follow the actions immediately; thus, animals must learn instrumental action–outcome contingencies using delayed reinforcement. Delays can hamper both Pavlovian and instrumental conditioning (Dickinson, 1980; Mackintosh, 1983; Dickinson, 1994; Gallistel, 1994; Hall, 1994): for example, although animals can bridge substantial delays to acquire instrumental responses, instrumental conditioning has long been ob-

served to be systematically impaired as the outcome is delayed (Skinner, 1938; Perin, 1943; Grice, 1948; Harker, 1956; Lattal & Gleeson, 1990; Dickinson *et al.*, 1992). Furthermore, individual variation in the ability to use delayed reinforcement may determine one aspect of impulsivity: an animal able to forgo short-term poor rewards in order to obtain delayed but better rewards may be termed self-controlled, whereas an animal that cannot tolerate delays to reward may be said to exhibit impulsive choice (Ainslie, 1975; Evenden, 1999b; Evenden, 1999a; Ainslie, 2001).

There are several psychological reasons why action–outcome delays might impair learning or performance of an instrumental response (Ainslie, 1975; Cardinal *et al.*, 2004). Instrumental responding is controlled by several processes (Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a); for example, rats work for outcomes that they value, using knowledge of the action–outcome contingencies in force to produce goal-directed actions. They also develop direct stimulus–response (S–R) associations, or habits. Action–outcome delays might, therefore, reduce the instrumental incentive value of the goal: valuing the goal less, animals may work less for it. Similarly, delays may hinder animals' ability to perceive the action–outcome contingency. Delayed rewards may also be less effective at reinforcing S–R habits. It is presently not known whether responses acquired with delayed reinforcement are governed by a different balance of habits and goal-directed actions than responses acquired with immediate reinforcement. However, one important factor in learning to act using delayed reinforcement may be the role of the environmental context. The animal's task is to attribute the outcome to its actions; instead, it may erroneously associate the outcome with the context, since the context is a cue that is temporally closer to the outcome than the action. The longer the delay, the more this contextual competition comes to impair the learning of the action–outcome contingency. Instrumental conditioning with delayed reinforcement can be enhanced if rats are exposed to the relevant contextual cues prior to instrumental training, and this enhancement is lessened if "free" (non-contingent) rewards are given during the contextual pre-exposure periods (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994). These results are consistent with the theory that during the action–outcome delay, contextual cues compete with the action to become associated with the outcome; pre-exposing the animals to the context with no consequences reduces this contextual competition, by making the context a bad predictor of the outcome (perhaps via latent inhibition or learned irrelevance), and this in turn makes the action–outcome contingency more salient and easier to learn (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994).

Little is known of the neural basis of instrumental learning with delayed reinforcement (Cardinal *et al.*, 2004). However, there is good evidence that the hippocampus contributes to the representation of context. Lesions of the hippocampal formation (H) have been shown to impair Pavlovian conditioning to a contextual CS, but not to a discrete CS, in rats (Hirsh, 1974; Selden *et al.*, 1991; Kim & Fanselow, 1992; Phillips & LeDoux, 1992; Honey & Good, 1993; Jarrard, 1993; Kim *et al.*, 1993; Phillips & LeDoux, 1994; Phillips & LeDoux, 1995; Chen *et al.*, 1996; Maren & Fanselow, 1997; Anagnostaras *et al.*, 1999; Rudy *et al.*, 2002), at least for some processes involving contextual representation (Good & Honey, 1991; Holland & Bouton, 1999; Good, 2002). Since context–outcome associations are thought to hinder instrumental learning with delayed reinforcement (contextual competition) (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994), it follows that if H lesions impair the formation of associations involving the context, such lesions might reduce contextual competition and hence *facilitate* instrumental conditioning when there is an action–outcome delay.

To investigate whether the hippocampus contributes to learning with delayed reinforcement, the present experiments examined the ability of rats with excitotoxic lesions of the hippocampus to acquire instrumental responding with delayed reward, comparing them to sham-operated controls. Each subject was

allowed to respond freely on two levers, one of which produced reinforcement after a delay of 0, 10, or 20 s (Figure 22, p. 53). H-lesioned rats were slightly impaired at learning the lever-press response in the absence of delays. Delays retarded learning in sham-operated controls, but the delays did not impair the H-lesioned rats to the same extent. Thus, as the delays were increased, H-lesioned rats became better at learning relative to controls, suggesting that the presence of delays had less of an effect on H-lesioned rats. To establish whether this relative improvement in learning with delayed reinforcement would also manifest itself as improved self-control, a different group of rats were also trained on a task in which they had to choose between an immediate, small reward and a delayed, large reward (Figure 15, p. 38) and made excitotoxic hippocampal lesions, before retesting the rats postoperatively. Good learning with delayed reinforcement did not translate to self-controlled choice. H lesions severely impaired rats' ability to choose the larger reward when it was delayed, but not when the delay preceding delivery of the large reward was removed, demonstrating that hippocampal lesions induce impulsive choice.

3.3 METHODS

3.3.1 Overview of experiments

3.3.1.1 Experiment 1: Effects of hippocampal lesions on acquisition of instrumental responding with delayed reinforcement

Forty-eight rats received excitotoxic lesions of the hippocampus ($n = 32$) or sham lesions ($n = 16$). Five died postoperatively. Subject were next trained in a task in which they had continuous access to two identical levers; one lever delivered a single food pellet each time it was pressed, and the other lever had no effect. For some rats, the food pellet was delivered immediately after the lever press (0 s condition; 9 H-lesioned rats and 5 shams). For others, each pellet was delayed by either 10 s (9 H, 6 sham) or 20 s (9 H, 5 sham). Subjects were trained for 14 sessions. They then had their locomotor activity assessed, and finally they were killed and perfused for histology.

3.3.1.2 Experiment 2: Effects of hippocampal lesions on choice involving delayed reinforcement

Twenty-four naïve rats were first trained to press levers for food and to nosepoke to initiate lever presentations in discrete trials. Subjects were then trained on a choice-of-delayed-reinforcement task (described below) for 19 sessions. After this, they were assigned to matched groups (as described below) to receive lesions of the hippocampus (H, $n = 16$) or sham lesions (sham, $n = 8$). Following recovery, they were retested on the basic task for 7 sessions to obtain a baseline measure of performance. After this, 4 sessions were given in which all delays were omitted in alternate sessions (DNDN design; D = delays present, N = no delays). Half of the subjects began this test with the delays present, and half with no delays (counterbalanced across groups). As a deficit was observed during testing (before histological data were available), further behavioural tests were given to elucidate the nature of the deficit. All subjects were given a further 6 sessions with no delays, in an attempt to re-equalize the two groups' performance and ensure that all animals would come to prefer the lever producing the large reinforcer. Delays were then re-introduced for a further 6 sessions. All subjects then underwent a food consumption test and had their locomotor activity assessed; finally, they were killed and perfused for histology.

3.3.2 Subjects and housing conditions

Subjects were male Lister hooded rats (Harlan-Olac UK Ltd) housed in a temperature-controlled room (minimum 22°C) under a 12:12 h reversed light–dark cycle (lights off 07:30 to 19:30). Subjects were approximately 15 weeks old on arrival at the laboratory and were given a minimum of a week to acclimatize, with free access to food, before experiments began. Experiments took place between 09:00 and 21:00, with individual subjects being tested at a consistent time of day. Subjects had free access to water, and were housed either in groups of four (Experiment 1) or in

pairs (Experiment 2). During behavioural testing, they were maintained at 85–90% of their free-feeding mass using a restricted feeding regimen. Feeding occurred in the home cages at the end of the experimental day. All procedures were subject to UK Home Office approval (Project Licence 80/1767) under the Animals (Scientific Procedures) Act 1986.

3.3.3 Excitotoxic lesions of the hippocampus

Subjects were anaesthetized with Avertin (2% w/v 2,2,2-tribromoethanol, 1% w/v 2-methylbutan-2-ol, and 8% v/v ethanol in PBS, sterilized by filtration, 10 ml/kg intraperitoneally) and placed in a Kopf or Stoelting stereotaxic frame (David Kopf Instruments, Tujunga, California, USA; Stoelting Co., Wood Dale, Illinois, USA) fitted with atraumatic ear bars. The skull was exposed and a dental drill was used to remove the bone directly above the injection and cannulation sites. The dura mater was broken with the tip of a hypodermic needle, avoiding damage to underlying venous sinuses. Excitotoxic hippocampal lesions targeted both the dorsal hippocampus and the ventral hippocampus. Lesions were made by injecting 0.09 M *N*-methyl-D-aspartic acid (NMDA; Sigma, UK) (Jarrard & Meldrum, 1993) through a glass micropipette (tip diameter 50–100 μm), using the coordinates, volumes, and timings shown in Table 5. The toxin had been dissolved in 0.1 M phosphate buffer (composition 0.07 M Na_2HPO_4, 0.028 M NaH_2PO_4 in double-distilled water, sterilized by filtration) and adjusted with NaOH to a final pH of 7.2–7.4. Sham lesions were made in the same manner except that vehicle was infused. At the end of the operation, animals were given 15 ml/kg of sterile 5% w/v glucose, 0.9% w/v sodium chloride intraperitoneally. Lesioned animals were given 0.2 ml of 5 mg/ml diazepam (Roche Products Ltd, UK) i.m. to prevent seizures. They were given two weeks to recover, with free access to food, and were handled regularly. Any instances of postoperative constipation were treated with liquid paraffin orally and rectally. At the end of this period, food restriction commenced or was resumed.

Region within hippocampus	Sites per hemisphere	AP	ML	DV	Volume injected per site	Duration of each infusion	Time allowed for diffusion after each infusion
Dorsal	2	−2.8	±1.6	−3.3	0.4 μl	4 min	3 min
		−4.2	±2.6	−3.0	0.4 μl	4 min	3 min
Ventral	4	−4.8	±4.8	−6.0	0.2 μl	2 min	3 min
		−5.3	±4.6	−4.2	0.2 μl	2 min	3 min
		−5.3	±4.6	−6.0	0.2 μl	2 min	3 min
		−5.8	±4.6	−4.2	0.2 μl	2 min	3 min

Table 5: Parameters for excitotoxic hippocampal lesions
Excitotoxic lesions of the entire hippocampus were made by injecting 0.09 M NMDA at the coordinates shown (see Methods). Along the anteroposterior (AP), mediolateral (ML), and dorsoventral (DV) axes, positive coordinates are in the anterior, left, and superior directions respectively. All coordinates are in mm. DV coordinates are measured from the dura above the injection site.

3.3.4 Behavioural apparatus

Behavioural testing was conducted in one of two types of operant chamber of identical configuration (from Med Associates Inc., Georgia, Vermont, USA, or Paul Fray Ltd, Cambridge, UK). Each chamber was fitted with a 2.8 W overhead house light and two retractable levers on either side of an alcove fitted with an infrared photodiode to detect head entry and a 2.8 W lightbulb ("traylight"). Sucrose pellets (45 mg, Rodent Diet Formula P, Noyes, Lancaster, New Hampshire, USA) could be delivered into the alcove. The chambers were enclosed within sound-attenuating boxes fitted with fans to provide air circulation. The apparatus was controlled by software written by RNC in C++ (Stroustrup, 1986) using the Whisker control system (Cardinal, 2000; Cardinal & Aitken, 2001).

3.3.5 Instrumental conditioning with delayed reinforcement (Experiment 1)

A variety of free-operant schedules may be used to assess instrumental acquisition with delayed reinforcement (Dickinson *et al.*, 1992). The simplest possible free-operant schedule was used in the present experiments (Cardinal & Cheung, 2005) (Chapter 2): each response scheduled a reinforcer after the programmed delay (Figure 22, p. 53). In such a schedule, if the subject responds during the delay, the experienced response–reinforcer delay will not match the programmed delay (as the second response is temporally close to the first reinforcer). However, this schedule has the advantage that the response–reinforcer contingency is constant (every response does in fact cause the delivery of reinforcement) and the reinforcement rate is not constrained (Dickinson *et al.*, 1992). So that re-sponding could be attributed to the instrumental response–reinforcer contingency, rather than the effects of general activity or reinforcement itself, responding on the active lever was compared to responding on a control lever that had no programmed consequence. Different groups of lesioned and sham-operated subjects were trained using dif-ferent delays; the delay was consistent for every subject. Delays of 0, 10, and 20 s were used.

Immediately after subjects were placed in the operant chamber, the sessions began. The houselight was illumi-nated, and remained on for each 30-min session. Two levers were extended into the chamber. All lever responses were first debounced to 10 ms (i.e. if a response occurred within 10 ms of a previous valid response it was attributed to mechanical bounce and ignored). Other than this, all lever presses and nosepokes into the food alcove were re-corded. Responding on the left (active) lever caused a single pellet to be delivered following a delay, under an FR-1 schedule (Figure 22, p. 53). To attribute acquisition of a lever-press response to the instrumental contingency, it is also necessary to control for the effects of reinforcer delivery itself (Dickinson *et al.*, 1992); therefore, responding on the active lever was compared to responding on the right (inactive) lever, which had no programmed conse-quence. To minimize any potential contribution of conditioned reinforcement to the task, no explicit signals were associated with pellet delivery other than the noise of the pellet dispenser apparatus. The schedule was implemented in the SimpleSchedules program (Cardinal, 2002b).

3.3.6 Lever and nosepoke training prior to the delayed reinforcement choice task (Experiment 2)

Subjects were first trained under an FR-1 schedule (where every lever press leads to the immediate delivery of a pellet) with only one lever present, to a criterion of a total of 50 presses on that lever across 30-min sessions, first for the left lever and then for the right. Subjects were then trained on a simplified version of the full task. The session began with the levers retracted and the operant chamber in darkness. Trials began every 40 s with the illumination of the houselight and the traylight. The subject was required to make a nosepoke response within 10 s, or the trial was aborted and the chamber returned to darkness (scored as an omission). If the subject nosepoked within the time limit, the traylight was extinguished and a single lever was presented (left/right at random). Subjects were required to respond on the lever within 10 s or the lever was retracted and the chamber darkened (scored as an omission). Upon pressing the lever, the houselight was switched off, a single pellet was delivered immediately and the traylight was illuminated until either the pellet was collected or 10 s had elapsed, whereupon the chamber was darkened, and the trial was counted as successful. Rats were trained to a criterion of 60 successful trials in one hour (the maximum possible with trials lasting 40 s being 90). The schedule was implemented in the ImpulsiveChoice program (Cardinal, 2002a).

3.3.7 Choice between small, immediate and large, delayed rewards (Experiment 2)

The task was based on Evenden & Ryan's (1996) procedure and has been described before (Cardinal *et al.*, 2000; 2001; Winstanley *et al.*, 2004b). The session began in darkness with the levers retracted; this was designated the intertrial state. Trials began at 100-s intervals; the format of a single trial is shown in Figure 15 (p. 38). Each trial began with the illumination of the houselight and the traylight. The rat was required to make a nosepoke response, ensuring that it was centrally located at the start of the trial (latency to poke was designated the initiation latency). If the rat did not respond within 10 s of the start of the trial, the operant chamber was reset to the intertrial state until

the next trial began and the trial was scored as an omission. If the rat was already nosepoking when the trial began, the next stage followed immediately. Upon a successful nosepoke, the traylight was extinguished and one or both levers were extended. One lever was designated the Delayed lever, the other the Immediate lever (counterbalanced left/right). The latency to choose a lever was recorded. (If the rat did not respond within 10 s of lever presentation, the chamber was reset to the intertrial state until the next trial and the trial was scored as an omission.) When a lever was chosen, both levers were retracted and the houselight was switched off. Choice of the Immediate lever caused the immediate delivery of one pellet; choice of the Delayed lever caused the delivery of 4 pellets following a delay. When reinforcement was delivered, the traylight was switched on. Multiple pellets were delivered 0.5 s apart. If the rat collected the pellets before the next trial began, then the traylight was switched off and time from delivery of the first pellet until a nosepoke occurred was recorded as the collection latency. If the rat did not collect the food within 10 s of its delivery, the operant chamber entered the intertrial state, though collection latencies were still recorded up to the start of the next trial. The chamber was then in the intertrial state and remained so until the next trial. There was no mechanism to remove uneaten pellets, but failure to collect the reward was an extremely rare event.

The delay was varied systematically across the session. A session 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 free-choice trials. Preferences were calculated for each block from only those trials on which the subject responded. Delays for each block were 0, 10, 20, 40 and 60 s respectively. As trials began every 100 s, the total session length was 100 minutes; subjects received one session per day. The schedule was implemented in the ImpulsiveChoice program (Cardinal, 2002a).

Preoperatively, subjects were trained on this task for 19 sessions. To allocate subjects into matched groups for surgery, their degree of sensitivity to the effects of delays within each session was assessed for the last three preoperative sessions, by calculating the slope of the linear regression of percentage choice of the large delayed reinforcer against delay for each subject. Rats were ranked by this measure, and rats with equivalent levels of performance were randomized to receive sham or H lesion surgery: the ranked list was divided into ordered triplets, and from each triplet one subject was assigned to the sham group and the other two to the H group, at random.

3.3.8 Locomotor activity in a novel environment

Locomotor activity was measured in wire mesh cages, 25 (W) × 40 (D) × 18 (H) cm, equipped with water bottles and two horizontal infrared photocell beams situated 1 cm from the floor that enabled movements along the long axis of the cage to be registered. The apparatus was controlled by software written by RNC in Arachnid (Paul Fray Ltd, Cambridge), a real-time extension to BBC BASIC V running on an Acorn Archimedes series computer. Subjects were placed in these cages, which were initially unfamiliar to them, and their activity was recorded for 2 h. All animals were tested in the food-deprived state.

3.3.9 Food consumption tests

Food consumption was assessed using four tests, conducted in subjects' home cages (always with only one rat present) on separate days under conditions of food deprivation. (1) Subjects were given free access to the 45-mg sucrose pellets used as reinforcers (Rodent Diet Formula P, Noyes, Lancaster, New Hampshire, USA) for 30 minutes; the amount eaten was recorded. (2) This test was repeated with the chow used as the maintenance diet. (3) The time taken to consume 50 sucrose pellets was recorded. (4) The time taken to consume an equivalent mass of chow (2.25 g) was recorded.

3.3.10 Histology

Rats were deeply anaesthetized with pentobarbitone sodium (200 mg/ml, minimum of 1.5 ml i.p.) and perfused transcardially with 0.01 M PBS followed by 4% paraformaldehyde in PBS. Their brains were removed and postfixed in paraformaldehyde before being dehydrated in 20% sucrose for cryoprotection. The brains were sectioned coronally at 60 μm thickness on a freezing microtome and every third section mounted on chromium potassium sul-

phate/gelatin-coated glass microscope slides and allowed to dry. Sections were passed through a series of ethanol solutions of descending concentration (3 minutes in each of 100%, 95%, and 70% v/v ethanol in water) and stained for ~5 min with cresyl violet. This stain comprises 0.05% w/v aqueous cresyl violet (Raymond A. Lamb Ltd, Eastbourne, UK), 2 mM acetic acid, and 5 mM formic acid in water. Following staining, sections were rinsed in water and 70% ethanol before being differentiated in 95% ethanol. Finally, they were dehydrated and delipidated in 100% ethanol and Histoclear (National Diagnostics, UK) before being cover-slipped using DePeX mounting medium (BDH, UK) and allowed to dry. The sections were used to verify lesion placement and assess the extent of lesion-induced neuronal loss. Lesions were detectable as the absence of visible neurons (cell bodies of the order of 100 μm in diameter with a characteristic shape), often associated with a degree of tissue collapse (sometimes with consequent ventricular expansion when the lesion was adjacent to a ventricle) and gliosis (visible as the presence of smaller, densely staining cells).

3.3.11 Data analysis

Data collected by the chamber control programs were imported into a relational database (Microsoft Access 97) for case selection and analysed with SPSS 11. Figures were created with SigmaPlot 2001/v7 and Adobe Illustrator 8. All graphs show group means and error bars are ±1 standard error of the mean (SEM) unless otherwise stated. Count data (lever presses and locomotor activity counts), for which variance increases with the mean, were subjected to a square-root transformation prior to any analysis (Howell, 1997). Homogeneity of variance was verified using Levene's test (Levene, 1960). General linear models are described as *dependent variable = $A_2 \times B_{cov} \times (C_5 \times D_{cov} \times S$)* where A is a between-subjects factor with two levels, B is a between-subjects covariate, C is a within-subjects factor with five levels, and D is a within-subjects covariate; S denotes subjects in designs involving within-subjects factors (Keppel, 1982). For repeated measures analyses, Mauchly's test of sphericity of the covariance matrix was applied (Mauchly, 1940) and the degrees of freedom corrected to more conservative values using the Huynh–Feldt epsilon $\tilde{\varepsilon}$ for any terms involving factors in which the sphericity assumption was violated (Huynh & Feldt, 1970).

3.4 RESULTS

3.4.1 Histology

In Experiment 1, there were five postoperative deaths. No rats were excluded after histological analysis; final group sizes were 9 (H, 0 s delay)[14], 5 (sham, 0 s delay)[15], 9 (H, 10 s delay)[16], 6 (sham, 10 s delay)[17], 9 (H, 20 s delay)[18], and 5 (sham, 20 s delay)[19]. In Experiment 2, there was one postoperative death (H group), and one rat (sham group) fell ill five sessions after surgery and was killed. Histological analysis revealed that the lesions were incomplete or encroached significantly on neighbouring structures in 3 subjects. These subjects were excluded; final group sizes were therefore 7 (sham)[20] and 12 (H)[21].

A diagram of the rat hippocampus was shown in Figure 12 (p. 28). Lesions of the hippocampus encompassed much of the dorsal and ventral hippocampal pyramidal cell (cornu ammonis CA1–CA3) fields, the dentate gyrus, the subiculum, and the fimbriae. Neuronal loss and associated gliosis extended in an anteroposterior direction from approximately –0.8 mm to –7.8 mm relative to bregma (negative coordinates are posterior). Damage to the dorsal and ventral hippocampal commissure was occasionally

---

[14] Experiment 1, H, 0 s delay: final subjects T2, T3, T4, T5, T6, T7, T8, T9, T15 (*n* = 9).
[15] Experiment 1, sham, 0 s delay: final subjects T10, T11, T12, T13, T14 (*n* = 5).
[16] Experiment 1, H, 10 s delay: final subjects T17, T18, T19, T20, T21, T22, T23, T24, T32 (*n* = 9).
[17] Experiment 1, sham, 10 s delay: final subjects T25, T26, T27, T28, T29, T30 (*n* = 6).
[18] Experiment 1, H, 20 s delay: final subjects T16, T31, T35, T36, T37, T38, T40, T47, T48 (*n* = 9).
[19] Experiment 1, sham, 20 s delay: final subjects T41, T42, T43, T45, T46 (*n* = 5).
[20] Experiment 2, sham: final subjects S1, S2, S7, S10, S13, S15, S17 (*n* = 7).
[21] Experiment 2, H: final subjects S3, S5, S6, S9, S11, S18, S19, S20, S21, S22, S23, S24 (*n* = 12).

seen, but damage to the overlying cortex was minimal. Schematics of the lesions are shown in Figure 37, and photomicrographs of a representative lesion are shown in Figure 38.
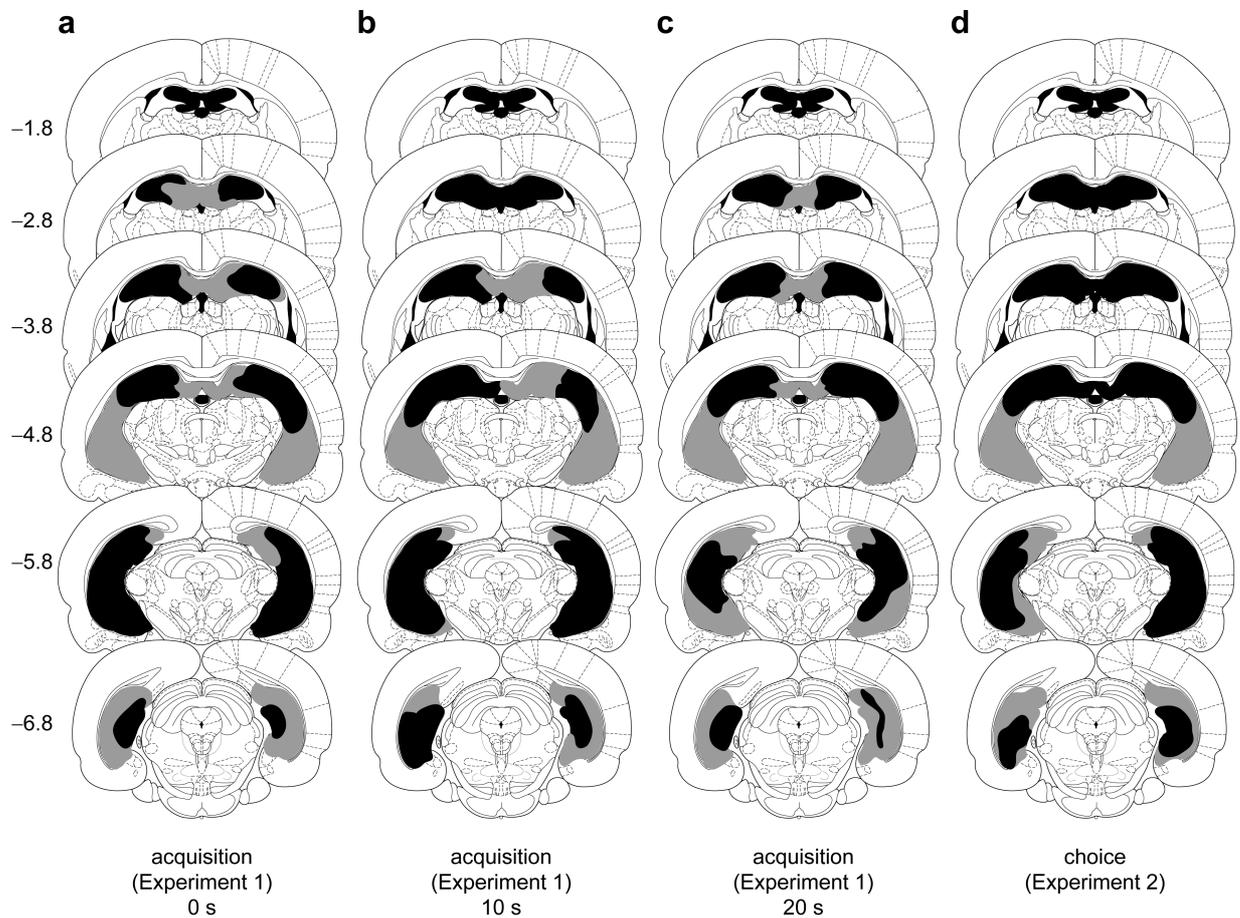
Figure 37: Schematic of lesions of the hippocampus

Black shading indicates the extent of neuronal loss common to all subjects (and also the third and lateral ventricles); grey indicates the area lesioned in at least one subject. Coronal sections are (from top to bottom) –1.8, –2.8, –3.8, –4.8, –5.8, and –6.8 mm relative to bregma. Diagrams are modified from Paxinos & Watson (1998). Panels **a–c** show schematics for Experiment 1 (acquisition of a free-operant instrumental response with delayed reinforcement; 0 s, 10 s, and 20 s groups, respectively) while **d** shows schematics for Experiment 2 (choice between small, immediate and large, delayed reinforcement).
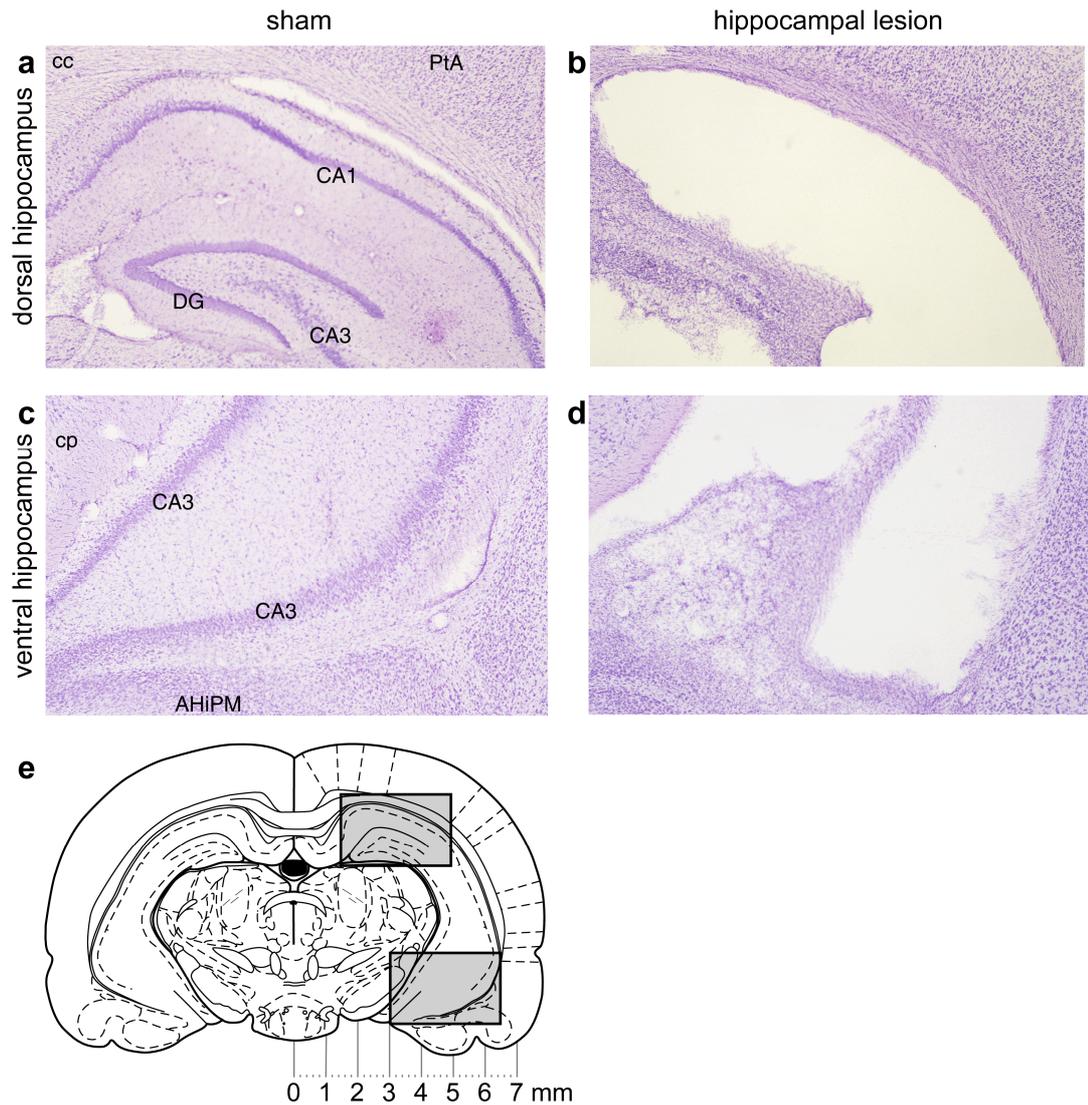
Figure 38: Photomicrographs of lesions of the hippocampus

Lesions of the hippocampus: photomicrographs of sections ~4.7 mm posterior to bregma, stained with cresyl violet. **(a)** Sham-operated rat, dorsal hippocampus, right hemisphere (medial to the left). CA1, cornu ammonis field 1; CA3, cornu ammonis field 3; DG, dentate gyrus; cc, corpus callosum; PtA, parietal association cortex. **(b)** H-lesioned rat; same area as (a). There is tissue collapse within the lesion and the ventricle is greatly expanded. **(c)** Sham-operated rat, ventral hippocampus. AHiPM, amygdalohippocampal area, posteromedial part; cp, cerebral peduncle. **(d)** H-lesioned rat, same area as (c). **(e)** Coronal diagram of the rat brain at 4.8 mm posterior to bregma (Paxinos & Watson, 1998), with scale. The upper grey box indicates approximately the region shown in (a) and (b); the lower grey box indicates approximately the region shown in (c) and (d).

3.4.2 Acquisition of instrumental responding (Experiment 1)

As expected, response–reinforcer delays retarded the acquisition of instrumental responding in sham-operated rats (Figure 39a). However, this impairment was lessened in H-lesioned rats (Figure 39b). H-lesioned rats responded less than shams in the absence of a response–reinforcer delay (Figure 40a), but responded as well as shams when delays were imposed (Figure 40b,c); H-lesioned rats were even facilitated numerically relative to shams in the 20 s delay condition (Figure 40c), though this difference was not statistically significant on its own. These conclusions were reached statistically as follows.

An overall ANOVA, using the model lesion$_2$ × delay$_3$ × (session$_{14}$ × lever$_2$ × S), revealed a lesion × lever × delay interaction ($F_{2,37} = 4.16$, $p = .023$), justifying sub-analyses, in addition to effects of delay ($F_{2,37} = 17.9$, $p < .001$), lever ($F_{1,37} = 435$, $p < .001$), delay × lever ($F_{2,37} = 4.16$, $p < .001$), session ($F_{5.35,198.0} = 38.7$, $\tilde{\varepsilon} = .412$, $p < .001$), delay × session ($F_{10.7,198.0} = 3.03$, $p = .001$), session × lever ($F_{4.99,184.6} = 17.5$, $\tilde{\varepsilon} = .384$, $p < .001$), and delay × session × lever ($F_{10.0,184.6} = 2.30$, $\tilde{\varepsilon} = .384$, $p = .015$). The differences between the groups were in their responding on the active lever (active lever, lesion × delay: $F_{2,37} = 3.71$, $p = .034$) rather than on the inactive lever (inactive lever, terms involving lesion: maximum $F_{2,37} = 1.146$, NS). All six groups learned to discriminate between the two levers, responding more on the active lever than on the inactive lever ($p < .05$, main effect of lever for each group).

Delays reduced the rate of acquisition and the final level of responding on the active lever for sham-operated rats (Figure 39a; delay, $F_{2,13} = 58.7$, $p < .001$; delay × session, $F_{10.8,70.4} = 2.67$, $\tilde{\varepsilon} = .417$, $p = .007$). Delays also increased responding on the inactive lever somewhat (Figure 39a; delay: $F_{2,13} = 5.26$, $p = .021$; delay × session, $F_{13.1,85.2} = 1.22$, $\tilde{\varepsilon} = .504$, NS). Similarly, in H-lesioned rats, delays reduced responding on the active lever (Figure 39b; delay: $F_{2,24} = 12.3$, $p < .001$; delay × session: $F_{7.8,93.1} = 2.76$, $\tilde{\varepsilon} = .298$, $p = .009$), although they did not significantly affect responding on the inactive lever (delay: $F_{2,24} = 1.91$, NS; delay × session: $F_{12.3,147.9} = 1.37$, $\tilde{\varepsilon} = .474$, NS).

At 0 s delay, H-lesioned rats responded significantly less than shams on the active lever (Figure 40a; lesion: $F_{1,12} = 6.11$, $p = .029$). There were no differences in responding on the inactive lever ($Fs < 1$, NS). At 10 s delay, there were no differences between sham-operated and H-lesioned rats in responding on either the active or the inactive lever ($Fs < 1.35$, $p \geq .266$). At 20 s delay, there were also no significant differences on either lever (active lever: lesion $F_{1,12} = 2.485$, $p = .141$, lesion × session $F < 1$, NS; inactive lever: $Fs < 1$, NS), although the H-lesioned rats responded numerically more than shams on the active lever throughout.

Inspection of Figure 39 also suggested that delays had less of an impact on the final (asymptotic) rates of responding in H-lesioned rats than in shams. The sessions were divided by eye into an "acquisition" phase (sessions 1–6) and a "stable" phase (sessions 7–14). Responding on the active lever in the stable phase was analysed; this revealed a lesion × delay interaction ($F_{2,37} = 3.44$, $p = .043$), with delays markedly reducing stable rates of responding in shams ($F_{2,13} = 42.3$, $p < .001$) but less so in H-lesioned rats ($F_{2,24} = 3.11$, $p = .063$).
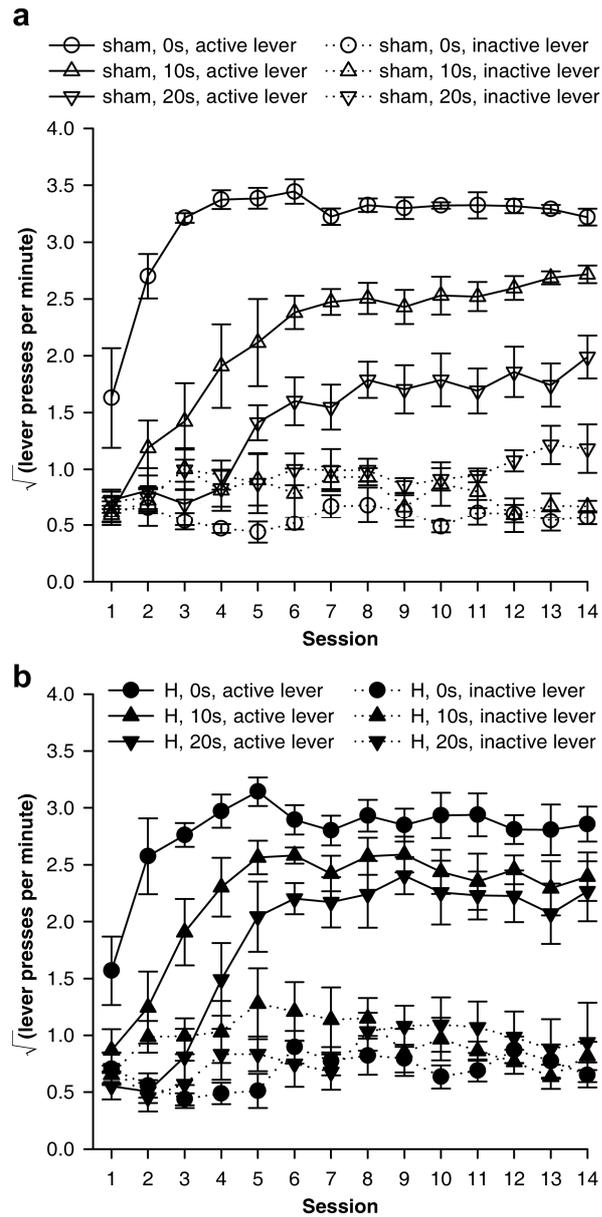
Figure 39: Effects of delays to reinforcement on acquisition of free-operant responding under an FR-1 schedule

Data plotted to show the effects of delays. All groups discriminated between the active and the inactive lever, and delays retarded acquisition of the active lever response in both groups. **(a)** Responding of sham-operated control rats, under all three response–reinforcer delay conditions. **(b)** Responding of H-lesioned rats under all delay conditions. The next figure replots these data to show the effect of the lesion more clearly.
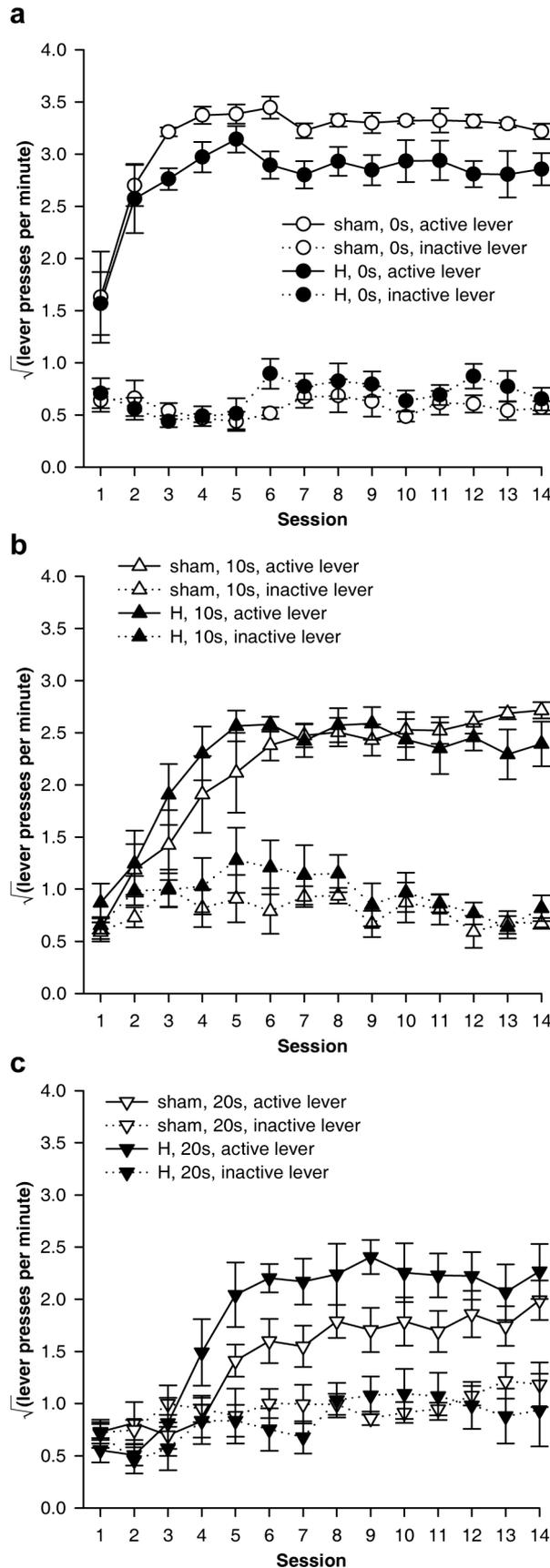
Figure 40: Effect of hippocampal (H) lesions on acquisition of free-operant responding with delayed reinforcement

Data plotted to show the effects of hippocampal lesions (same data as in the previous figure). There was a delay-dependent impairment in H-lesioned rats (significant lesion × delay interaction, see text), who learned less well than shams only when reinforcement was *not* delayed. **(a)** With a delay of 0 s, H-lesioned rats responded less on the active lever than shams did. **(b)** With a 10 s delay, H-lesioned rats responded the same as shams. **(c)** With a 20 s delay, H-lesioned rats responded more than shams on the active lever, though this difference was not statistically significant on its own.

3.4.3 Experienced response–delivery and response–collection delays (Experiment 1)

For every reinforcer delivered, the active lever response most closely preceding it in time was identified, and the time between that response and delivery of the reinforcer (the "response–delivery delay") was calculated. This time can therefore be equal to or less than the programmed delay, and is only relevant for subjects experiencing non-zero programmed response–reinforcer delays. The response-to-reinforcer-collection ("response–collection") delays were also calculated: for every reinforcer delivered, the response most closely preceding it and the nosepoke most closely following it were identified, and the time between these two events calculated. This time can be shorter or longer than the programmed delay, and is relevant for all subjects.

H-lesioned rats experienced slightly shorter response–delivery delays than shams when the programmed delay was 10 s or 20 s (Figure 41a): there was a lesion × programmed delay interaction ($F_{1,25} = 6.28$, $p = .019$), and simple effects of the lesion when the programmed delay was 10 s ($F_{1,13} = 8.49$, $p = .012$) and when it was 20 s ($F_{1,12} = 9.50$, $p = .009$).

H-lesioned rats also experienced slightly shorter response–collection delays across all programmed delays (Figure 41b) (lesion: $F_{1,37} = 4.21$, $p = .047$), though the difference was not significant at any one programmed delay (lesion × programmed delay: $F_{2,37} = 2.35$, $p = .109$; simple effects of the lesion at different programmed delays: maximum $F_{1,12} = 3.08$, $p = .105$).

These differences in the *mean* delay experienced by each rat were reflected in differences in the *distribution* of response–delivery and response–collection delays when the programmed delay was non-zero (Figure 41c,d). All experienced delays for a given subject were aggregated across all sessions, and the proportion falling into different 2-s ranges were calculated to give one value per range per subject. For response–delivery delays, H-lesioned rats experienced slightly fewer long delays and slightly more short delays in the 10 s condition (lesion × range, $F_{3.6,47.0} = 3.40$, $\tilde{\varepsilon} = .723$, $p = .019$) and in the 20 s condition (lesion × range, $F_{1.4,16.6} = 6.54$, $\tilde{\varepsilon} = .138$, $p = .014$). For response–collection delays, there were no differences in the distribution of delays experienced by H-lesioned and sham rats in the 0 s condition (lesion and lesion × range, Fs < 1, NS). In the 10 s condition, H-lesioned rats experienced a slightly lower proportion of long response–collection delays and a slightly higher proportion of short response–collection delays (lesion × range, $F_{4.4,57.6} = 3.60$, $\tilde{\varepsilon} = .233$, $p = .009$). Similarly, in the 20 s condition, H-lesioned rats experienced a slightly lower proportion of long response–collection delays and a slightly higher proportion of short response–collection delays than shams (lesion × range, $F_{3.1,37.5} = 7.02$, $\tilde{\varepsilon} = .164$, $p = .001$).

Since H-lesioned rats experienced slightly shorter delays than sham-operated rats, it was necessary to take this into account when establishing the effect of delays on learning, as follows.
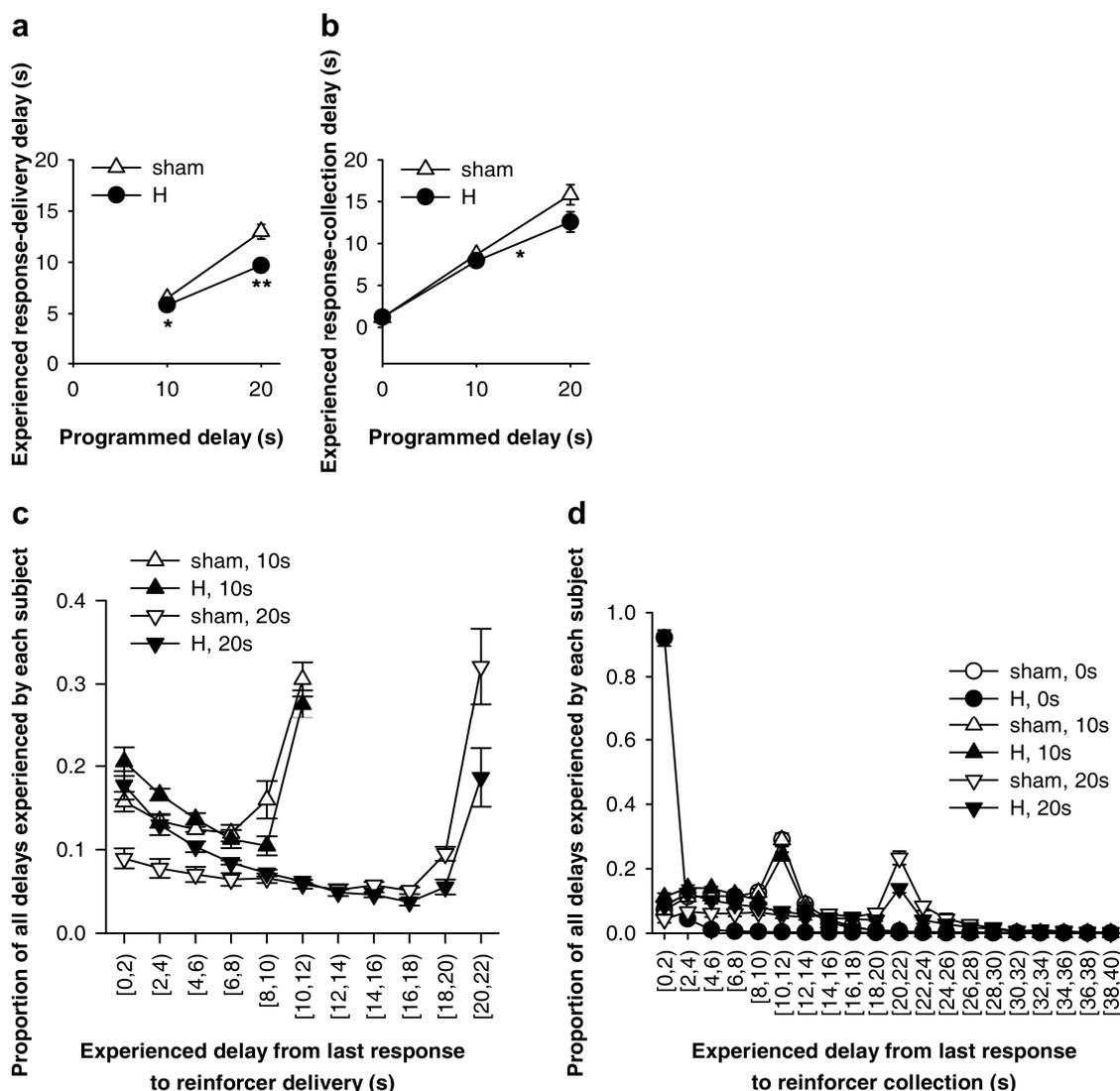
Figure 41: Programmed and experienced delays to reinforcement in H-lesioned and sham-operated rats

H-lesioned rats experienced slightly shorter response–delivery delays (the delay between the most recent lever press and pellet delivery) than shams, and slightly shorter response–collection delays (the delay between the most recent lever press and pellet collection). **(a)** Mean experienced response–delivery delays (one value calculated per subject). When the programmed delay was 0 s, reinforcers were delivered immediately so no data are shown. H-lesioned rats experienced shorter response–delivery delays when the programmed delay was 10 s (* p = .012) or 20 s (** p = .009). **(b)** Mean experienced response–collection delays (one value calculated per subject). H-lesioned rats experienced slightly shorter delays overall (* p = .047, main effect of lesion), but the experienced delays did not differ significantly at any given programmed delay. **(c)** Distribution of experienced response–delivery delays. All experienced delays for a given subject were aggregated across all sessions, and the proportion falling into different 2-s ranges were calculated to give one value per range per subject; the graphs show means ± SEMs of these values. The interval notation "[*a, b*)" indicates that a given delay *x* falls in the range $a \leq x < b$. H-lesioned rats experienced slightly fewer long delays and slightly more short delays in the 10 s condition (p = .019) and in the 20 s condition (p = .014). **(d)** Distribution of experienced response–collection delays, displayed in the same manner as (c). There were no differences in the distribution of delays experienced by H-lesioned and sham rats in the 0 s condition. In the 10 s condition, H-lesioned rats experienced a slightly lower proportion of long delays and a slightly higher proportion of short delays (p = .009), and similarly in the 20 s condition (p = .001).

3.4.4 Effect of delays on learning (Experiment 1)

There was a systematic relationship between the acquisition rate and the programmed delay of reinforcement, and this was altered in H-lesioned rats, who were less impaired by delays (compared to their performance at zero delay) than shams were. Figure 42a replots the rates of lever pressing on session 6, at the end of the initial "acquisition" phase. Despite the comparatively low power of such an analysis, lever pressing was analysed for this session only, using the model lesion$_2$ × delay$_3$. This revealed a significant lesion × delay interaction ($F_{2,37} = 8.67$, $p = .001$), which was analysed further. Increasing delays significantly reduced the rate of responding in this session for shams ($F_{2,13} = 31.4$, $p < .001$) and H-lesioned rats ($F_{2,24} = 8.88$, $p = .001$). H-lesioned rats responded less than shams at zero delay ($F_{1,12} = 8.08$, $p = .015$), were not significantly different from shams at 10 s delay ($F_{1,13} = 1.848$, $p = .197$), and responded more than shams at 20 s delay ($F_{1,12} = 6.23$, $p = .028$).

Since the H group experienced slightly shorter response–delivery and response–collection delays than shams when the programmed delay was non-zero (Figure 41), it is important to establish whether this effect alone was responsible for the lesser effect of delays on learning, or whether the effect of delays in H-lesioned rats was lessened over and above any effect to decrease the experienced delay. The mean experienced response–collection delay was calculated for each subject up to and including session 6. The square-root transformed number of lever presses in session 6 was then analysed using a general linear model of the form lesion$_2$ × experienced delay$_{cov}$ × S; unlike a standard analysis of covariance, the factor × covariate interaction term was included in the model. This confirmed that the detrimental effects of delay upon learning were reduced in H-lesioned rats, compared to controls, over and above the differences in experienced delay (Figure 42b; lesion × experienced delay: $F_{1,39} = 10.8$, $p = .002$).
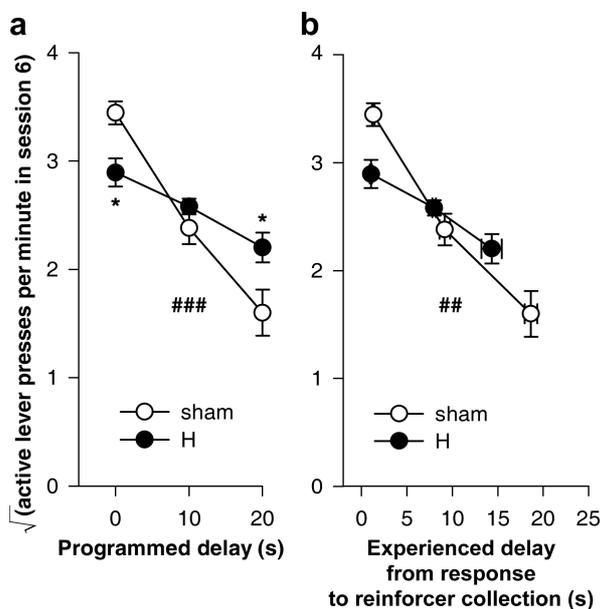
Figure 42: Learning as a function of programmed and experienced delays to reinforcement in H-lesioned and sham-operated rats

The imposition of response–reinforcer delays systematically retarded the acquisition of free-operant instrumental responding, but this effect was lessened in H-lesioned rats, even allowing for differences in experienced response–collection delays. **(a)** The rate of lever pressing in session 6 is plotted against the programmed response–reinforcer delay. There was a lesion × delay interaction (### $p = .001$): H-lesioned rats responded less than shams at zero delay (* $p = .015$), were not significantly different from shams at 10 s delay ($p = .197$), and responded more than shams at 20 s (* $p = .028$). **(b)** Responding in session 6 plotted against the experienced response-to-reinforcer collection delays for sessions 1–6 (vertical error bars: SEM of the square-root-transformed number of responses in session 6; horizontal error bars: SEM of the experienced response–collection delay, calculated up to and including that session). The gradients of the two lines differed significantly (## $p = .002$; see text), indicating that the relationship between experienced delays and responding was altered in H-lesioned rats.

3.4.5 Experienced delays and learning on the inactive lever (Experiment 1)

No such delay-dependent lesion effects were observed for the inactive lever. Experienced inactive-response–delivery delays (calculated across all sessions in the same manner as for the active lever) were

much longer and more variable than corresponding delays for the active lever, because subjects responded on the inactive lever so little. Means ± SEMs were 271 ± 31 s (sham, 0 s), 241 ± 23 s (H, 0 s), 201 ± 45 s (sham, 10 s), 184 ± 45 s (H, 10 s), 127 ± 21 s (sham, 20 s), and 171 ± 36 s (H, 20 s). ANOVA of these data showed that these experienced inactive-response–delivery delays depended upon the programmed active-response–delivery delay (delay: $F_{2,37} = 3.80$, $p = .032$) but there was no effect of the lesion and no lesion × delay interaction (Fs < 1, NS). Experienced inactive-response–collection delays were 272 ± 31 s (sham, 0 s), 242 ± 23 s (H, 0 s), 204 ± 45 s (sham, 10 s), 186 ± 45 s (H, 10 s), 130 ± 21 s (sham, 20 s), and 174 ± 35 s (H, 20 s). Again, ANOVA that these experienced delays depended upon the programmed active-response–delivery delays (delay: $F_{2,37} = 3.68$, $p = .035$) but there was no effect of the lesion and no lesion × delay interaction (Fs < 1, NS). When the square-root-transformed number of responses on the inactive lever in session 6 was analysed with the experienced delays up to that point as a predictor, using the model lesion$_2$ × experienced inactive-response–collection delay$_{cov}$ just as for the active lever analysis, there was no lesion × experienced delay interaction; neither was there an effect of lesion or experienced delay (maximum $F_{2,37} = 1.54$, NS).

3.4.6 Choice between an immediate, small reward and a large, delayed reward (Experiment 2)

Preoperatively, subjects preferred the larger reinforcer less when it was delayed, and the groups remained matched following later histological selection (Figure 43a). Choice ratios (percent choice of the large, delayed reinforcer, calculated for all free-choice trials on which subjects responded) from the last 3 preoperative sessions were analysed using the model lesion intent$_2$ × (delay$_5$ × S). While there was a strong effect of delay ($F_{2.2,36.9} = 22.9$, $p < .001$), no terms involving lesion intent were significant (Fs < 1, NS).

The choice patterns of the two groups diverged following surgery, with the H-lesioned rats choosing the large, delayed reinforcer less than sham-operated controls (Figure 43b). Comparison of choice in the last 3 preoperative sessions (Figure 43a) to that in the first 7 postoperative sessions (Figure 43b), using the model lesion$_2$ × (pre/post$_2$ × delay$_5$ × S) revealed a lesion × pre/post interaction ($F_{1,17} = 6.50$, $p = .021$). However, at this point, analysis of postoperative choice patterns on their own (Figure 43b) did not reveal a significant difference between the two groups (lesion: $F_{1,17} = 3.46$, $p = .08$; delay × lesion: $F < 1$, NS); as it did not take account of preoperative choice patterns, this analysis was less powerful. Later, H-lesioned rats diverged further from sham-operated controls and the difference between the two became significant even without taking account of preoperative information (see below). Both groups remained sensitive to the delay postoperatively (sham, effect of delay: $F_{2.1,12.8} = 4.59$, $\tilde{\varepsilon} = .531$, $p = .03$; lesion, effect of delay: $F_{1.5,16.5} = 7.05$, $\tilde{\varepsilon} = .374$, $p = .01$).

There were no differences between H-lesioned and sham-operated rats in any other measures collected, including the rate of omissions, the latency to initiate trials, the latency to choose a lever, the latency to collect food, and the rate of nosepoking in the food alcove during delays to reinforcement. Data from the 7 baseline postoperative sessions (sessions 20–26) were analysed. Omissions were very infrequent (overall, rats failed to initiate and/or press a lever on 0.2% of trials) and there were no group differences in the rates of omission ($F < 1$, NS). Initiation latencies did not differ between groups (lesion: $F < 1$, NS; lesion × delay, $F_{2.8,48.0} = 1.027$, $\tilde{\varepsilon} = .706$, NS). Neither did choice latencies: an analysis using the model lesion$_2$ × (delay$_5$ × lever$_2$ × S) revealed no significant terms involving lesion (Fs < 1.06, NS). Food collection latencies were analysed using the model lesion$_2$ × (choice$_2$ × delay$_5$ × S). Predictably, rats were slower to collect the food following choice of the large, delayed reinforcer as the delays got longer (choice × delay: $F_{2.4,38.5} = 19.8$, $\tilde{\varepsilon} = .602$, $p < .001$; effect of delay following choice of the small, immedi-

ate reinforcer: $F_{2.5,39.7} = 1.63$, $\tilde{\varepsilon} = .62$, NS; effect of delay following choice of the large, delayed reinforcer: $F_{2.2,35.8} = 18.4$, $\tilde{\varepsilon} = .559$, $p < .001$) but this was not influenced by the lesion (terms involving lesion, maximum $F_{1,16} = 1.93$, NS). The proportion of the delay spent nosepoking did not alter as a function of the delay, and was not affected by the lesion (only applicable to trials on which the large reinforcer was chosen with a non-zero delay; delay, $F_{1.3,22.3} = 2.38$, $\tilde{\varepsilon} = .437$, $p = .131$; lesion × delay, $F_{1.3,22.3} = 1.53$, $\tilde{\varepsilon} = .437$, NS; lesion: $F < 1$, NS).
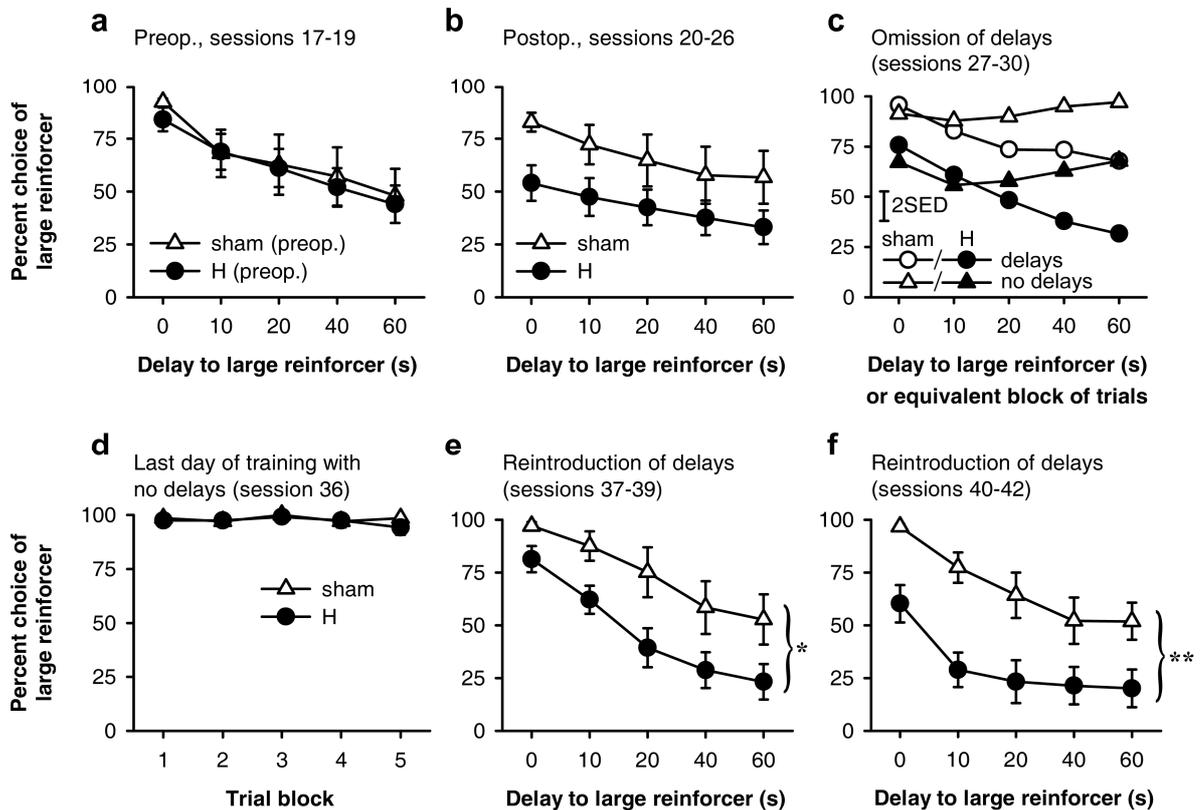
Figure 43: Effects of hippocampal lesions on choice between immediate, small rewards and large, delayed rewards

(a) Pattern of choice in the last three sessions before surgery; the sham and lesion groups were matched for performance. Rats' preference for the large reinforcer declined with delay ($p < .001$). **(b)** Choice in the first seven postoperative sessions. Although there was a change in behaviour in the lesioned group (lesion × pre/postop., $p = .021$), the difference between the two groups was not significant in its own right for these sessions ($p = .08$). **(c)** Effects of omitting all delays in alternating sessions (error bar, 2 SED for the three-way interaction). H-lesioned rats remained sensitive to the contingencies, altering their behaviour in response to delay omission, as shams did. **(d)** Last of six further consecutive sessions in which delays were omitted. Both groups preferred the large reinforcer strongly when it was not delayed, with no differences between sham and H-lesioned rats. **(e)** First three sessions following reintroduction of delays. H-lesioned rats were impulsive, choosing the large, delayed reinforcer less often than shams (* p = .027). **(f)** Next three sessions following reintroduction of delays. H-lesioned rats remained impulsive (** p = .007), and generalization between trial blocks occurred, reducing their preference for the large reinforcer in the zero-delay block as well (see text).

3.4.7 Effects of removing and reintroducing delays to the large reinforcer (Experiment 2)

Both H-lesioned and sham-operated rats were sensitive to the removal of delays in alternating sessions, increasing their preference for the large reinforcer during sessions when it was not delayed (Figure 43c). Choice ratios from these sessions were analysed using the model lesion$_2$ × (delays/no delays$_2$ × trial

$block_5 \times S$). This revealed a delays/no delays × block interaction ($F_{2.5,42.6} = 15.3$, $\tilde{\varepsilon} = .626$, $p < .001$). Additionally, there was a main effect of lesion ($F_{1,17} = 7.23$, $p = .016$), indicating a greater overall preference for the smaller reinforcer in H-lesioned rats compared to controls across these sessions, but there were no other significant terms involving lesion ($Fs < 1.05$, NS). In sessions when delays were present, both H-lesioned and sham-operated rats showed a within-session shift in preference as the delay increased (sham, effect of delay: $F_{4,24} = 4.79$, $p = .006$; H-lesioned, effect of delay: $F_{2.6,28.1} = 20.21$, $\tilde{\varepsilon} = .638$, $p < .001$), and H-lesioned rats chose the smaller, immediate reinforcer more often (lesion: $F_{1,17} = 5.91$, $p = .026$). In sessions when delays were not present, neither group showed a within-session shift in preference (shams: $F_{4,24} = 2.46$, $p = .073$; H-lesioned: $F_{2.1,23.5} = 1.63$, $\tilde{\varepsilon} = .534$, NS), though again the H-lesioned rats showed a stronger preference for the smaller reinforcer ($F_{1,17} = 6.61$, $p = .02$).

These analyses suggested that the H-lesioned rats' preference for the larger reward was less than that of shams even when it was not delayed. However, an alternative possibility is that the H-lesioned rats were loath to choose the large reinforcer when it was delayed, and that this generalized to affect preference even when it was not delayed (Cardinal *et al.*, 2001; 2003b). Consequently, subjects were given a further six sessions with no delays present; preference on the last of these sessions is shown in Figure 43d. H-lesioned rats showed a strong preference for the large reinforcer when it was not delayed, just as shams did, with mean choice ratios >94% in all conditions. In this session, there were no group differences ($Fs < 1$, NS) and no within-session shift in preference ($Fs < 1$, overall and for H-lesioned and sham-operated groups individually).

When delays were reintroduced (sessions 37–42), preference for the larger, delayed reinforcer declined much more sharply in H-lesioned rats than in shams (Figure 43e,f). Preference for the large reinforcer declined first at long delays, then progressively at shorter delays, such that even responding in the zero-delay block was affected. In sessions 37–39, H-lesioned rats chose the large reinforcer less often than shams (lesion: $F_{1,17} = 5.90$, $p = .027$; lesion × delay: $F_{1.9,32.7} = 1.16$, $\tilde{\varepsilon} = .482$, NS), with this difference being significant for 10 s and 20 s delays ($p < .05$) but not 0 s ($p = .075$), 40 s ($p = .058$), or 60 s delays ($p = .054$). In sessions 40–42, the pattern was essentially the same (lesion: $F_{1,17} = 9.47$, $p = .007$; lesion × delay: $F_{1.6,28.0} = 1.286$, $\tilde{\varepsilon} = .412$, NS), except that individual differences were now significant at all delays ($p < .05$).

3.4.8 Locomotor activity, body mass, and food consumption

H-lesioned animals were hyperactive compared to sham-operated controls in both experiments (Figure 44a,b), as reported previously (Good & Honey, 1997; McNish *et al.*, 1997). In Experiment 1, analysis of the square-root-transformed number of infrared beam breaks using the model $lesion_2 \times (bin_{12} \times S)$ revealed effects of lesion ($F_{1,41} = 9.77$, $p = .003$), reflecting hyperactivity in the H group, with additional effects of bin ($F_{8.2,335.7} = 58.4$, $\tilde{\varepsilon} = .744$, $p < .001$), reflecting habituation, and a lesion × bin interaction ($F_{8.2,335.7} = 2.95$, $\tilde{\varepsilon} = .744$, $p = .003$). In Experiment 2, hyperactivity was again observed (lesion: $F_{1,17} = 24.1$, $p < .001$; bin: $F_{8.6,145.7} = 15.9$, $\tilde{\varepsilon} = .779$, $p < .001$; lesion × bin: $F < 1$, NS).

In Experiment 1, they remained the same weight as sham-operated controls throughout, though in Experiment 2, which lasted longer, they gained less weight than shams (Figure 44c). There were no differences between groups preoperatively in either experiment ($Fs \leq 1.35$, NS); in Experiment 1, the groups gained weight at the same rate (lesion × time, $F < 1$, NS; group difference at second time point: $F < 1$, NS), but in Experiment 2, which lasted longer, the H-lesioned rats weighed less at the end of the experiment (lesion × time, $F_{1,15} = 14.5$, $p = .002$; group difference at second time point: $F_{1,15} = 8.56$, $p = .01$). Data from two H-lesioned subjects in Experiment 2 were lost.

H-lesioned rats consumed their maintenance chow more quickly and consumed more of it, but they did not differ from sham-operated controls in their consumption of the sucrose pellets employed as reinforcers in the behavioural tasks. In 30 minutes, H-lesioned rats consumed more chow (11.2 ± 0.6 g) than shams (8.2 ± 0.8 g) ($F_{1,7} = 8.36$, $p = .01$). However, there were no differences between the mass of sucrose pellets consumed in 30 minutes by H-lesioned rats (17.5 ± 1.2 g) and by shams (18.3 ± 1.6 g) ($F < 1$, NS). H-lesioned rats were quicker to consume 2.5 g of chow (taking 302 ± 18 s) than shams (385 ± 24 s) (Levene's test indicated significant heterogeneity of variance; Mann–Whitney $U_{7,12} = 18$, $p = .045$). However, although H-lesioned rats were also slightly quicker to consume 2.5 g of sucrose pellets (taking 160 ± 9 s) than shams (who took 176 ± 12 s), this difference was not significant ($F_{1,17} = 1.24$, $p = .281$).
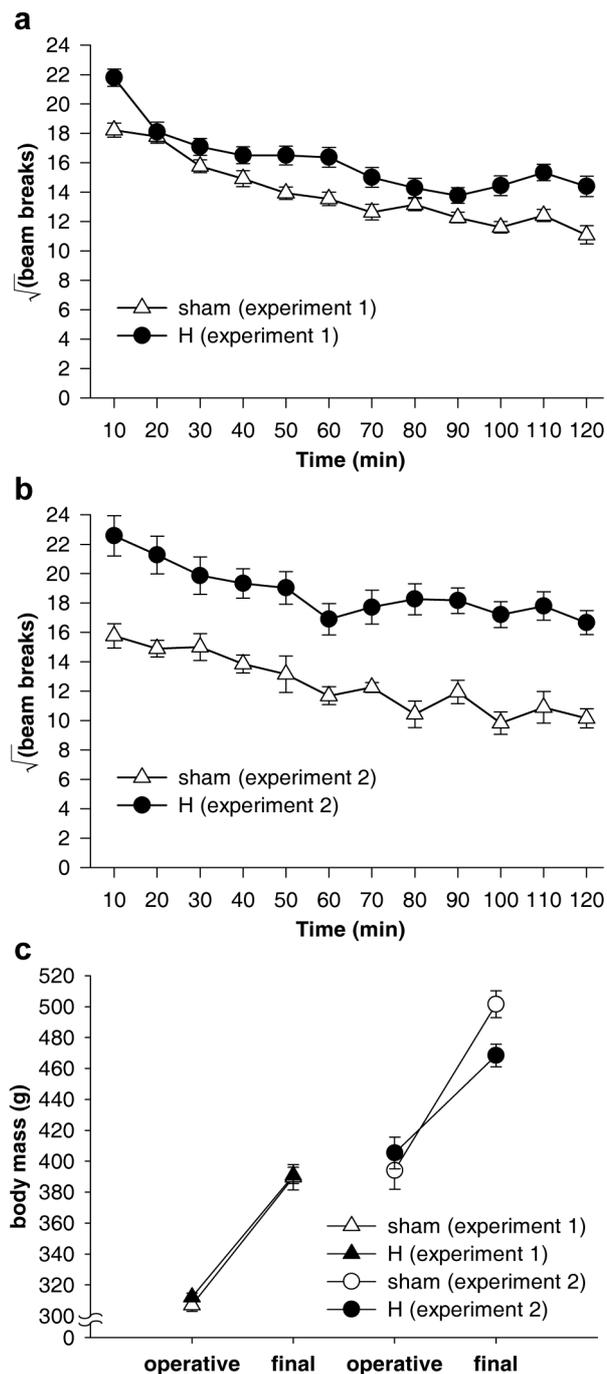
Figure 44: Locomotor activity in a novel environment and body mass in H-lesioned and sham-operated rats

H-lesioned rats were significantly hyperactive compared to sham-operated controls, in both **(a)** Experiment 1 ($p = .003$) and **(b)** Experiment 2 ($p < .001$). **(c)** Body mass across both experiments. There were no differences between groups preoperatively in either experiment; in Experiment 1, the groups gained weight at the same rate, but in Experiment 2, which lasted longer, the H-lesioned rats weighed less at the end of the experiment ($p = .01$).

3.5 DISCUSSION

Excitotoxic lesions of the dorsal and ventral hippocampus slightly retarded instrumental learning on a continuous reinforcement (FR-1) schedule in the absence of response–reinforcer delays. However, H-lesioned rats were only impaired when reinforcement was delivered immediately, and not when it was delivered after a delay. H-lesioned rats were less sensitive to the deleterious effects of response–reinforcer delays on learning, to the extent that with long (20 s) response–reinforcer delays, H-lesioned rats showed numerically better discrimination between the active and inactive levers than shams (Figure 40, Figure 42). Despite this delay-dependent facilitation of instrumental conditioning, H-lesioned rats were less able than shams to choose a delayed, large reinforcer in preference to an immediate, small reinforcer (Figure 43). That is, H-lesioned rats exhibited impulsive choice.

3.5.1 Pavlovian and instrumental conditioning with delayed reinforcement

Free-operant instrumental conditioning, and instrumental discrimination learning, have long been known to be impaired systematically by response–reinforcer (action–outcome) delays (Skinner, 1938; Grice, 1948; Lattal & Gleeson, 1990; Dickinson *et al.*, 1992). This might be for several reasons (Mackintosh, 1974; Ainslie, 1975; Cardinal *et al.*, 2004) because instrumental responding depends on several processes, including knowledge of the action–outcome contingency, a representation of the instrumental incentive value of the outcome, S–R habits, and the influence of Pavlovian CSs that have motivational significance through processes such as conditioned reinforcement and PIT (Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a). Action–outcome delays might affect several of these processes. For example, such delays may hinder the subject's ability to perceive the action–outcome contingency, so the subject is unaware that its actions will result in the outcome. Delays might reduce the value of the goal, so the subject is less willing to work for it. Delays might also impair the process by which S–R habits are reinforced; finally, they might affect the degree to which stimuli associated with reinforcement by Pavlovian conditioning are capable of motivating behaviour. In the present experiment, no explicit stimuli were presented associated with either the response or the reinforcer, to minimize the possible contribution of processes such as conditioned reinforcement. Nevertheless, it is largely an open question which processes contributing to instrumental learning and performance are the ones most affected by response–reinforcer delays; for example, it is not present known whether responses acquired with delayed reinforcement are governed by a different balance of habits and goal-directed actions than responses acquired with immediate reinforcement.

However, one process has been clearly demonstrated to influence learning with delays, and that process involves the environmental context. The effect of contextual factors on learning was first demonstrated in Pavlovian conditioning; the two paradigms typically used are known as *delay conditioning* and *trace conditioning*. In both cases, the interstimulus interval (ISI), which is the time between the onset of the CS and the onset of the US, is greater than zero. When a CS is followed by a US and the two *overlap* and are contiguous in time, the paradigm is known as "delay" conditioning (the terminology is somewhat confusing). When a CS is followed by a US and the two do *not* overlap, so that there is a gap between the end of the CS and the start of the US, the paradigm is known as "trace" conditioning (because the US must be associated with a "trace" of the CS) (Mackintosh, 1974). "Trace" Pavlovian conditioning, with a CS–US gap, results in poorer learning than "delay" Pavlovian conditioning, even if the ISI is held constant (Pavlov, 1927; Ellison, 1964; Kamin, 1965; Schneiderman, 1966; Mackintosh, 1974; McEchron *et al.*, 1998; McEchron & Disterhoft, 1999; Weiss *et al.*, 1999b; Beylin *et al.*, 2001; Desmedt *et al.*, 2003), just as insertion of an action–outcome gap retards instrumental conditioning (Grice, 1948; Lattal & Glee-

son, 1990; Dickinson *et al.*, 1992), although delay conditioning can, if necessary, be made as hard as trace conditioning for purposes of direct comparison by extending the ISI in the delay condition (Ivkovich *et al.*, 2000; Beylin *et al.*, 2001). It has been suggested that delay and trace conditioning proceed via different psychological mechanisms in humans, with the former depending on procedural memory and the latter on declarative memory (Clark *et al.*, 2001a). Regardless, several lines of evidence suggest that contextual competition might reduce responding to the discrete CS in trace conditioning. As the trace gap is lengthened, CRs tend to occur to the context instead of to the discrete CS (Odling-Smee, 1975; Marlin, 1980; Marlin & Miller, 1981; Selden *et al.*, 1991; Parkinson *et al.*, 1999b; Desmedt *et al.*, 2003). Trace conditioning can be improved by the addition of a "filler" stimulus during the CS–US gap, which might decrease contextual competition or act as a secondary or conditioned reinforcer (Kamin, 1965; Kaplan & Hearst, 1982; Rescorla, 1982). The smaller the ratio of the ISI to the intertrial interval (ITI; the time between the end of the US and the start of the next trial's CS), the faster conditioning proceeds (Gallistel, 1994), and one explanation of this is that long ITIs reduce the strength of context–US associations, making CS–US associations more salient. Finally, pre-exposure to the context in the absence of any US improves subsequent conditioning to a discrete CS, as would be expected under the contextual competition account since pre-exposure should produce latent inhibition of the context (see Boughner *et al.*, 2004).

The idea that reinforcing outcomes may be associated with either a discrete predictor (such as a Pavlovian CS or an instrumental response) or the background context, and that the two compete in some way for such association, also explains observations concerning the effect of contextual manipulations on instrumental conditioning (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994). Dickinson *et al.* (1992) trained rats on a free-operant, FR-1 schedule of reinforcement very similar to the one used in the present experiments, except that responding on the active lever was compared not to responding on an inactive lever, but to responding of a yoked control group (who received the same pattern of reinforcement as the "master" rats but whose lever presses had no consequence). They found that the rate of learning, and the asymptotic level of responding, declined across groups as the response–reinforcer delay was increased from 0 to 32 s; rats trained with a 64-s delay failed to learn at all, compared to yoked controls. However, when rats were exposed to the training context, in the absence of the lever or any reinforcers, prior to training, their learning was improved, and successful discrimination was seen even with a delay of 64 s. This is exactly what would be expected if a process of contextual competition was operating. The subject's task is to distinguish $P(outcome \mid action)$ from $P(outcome \mid no\ action)$, or, making the contribution of the context explicit, to distinguish $P(outcome \mid action + context)$ from $P(outcome \mid context)$. Pre-exposure to the context would be expected to produce latent inhibition to the context, reducing the strength of context–outcome associations. Viewed another way, non-reinforced exposure to the context forces the subjects to experience a zero-response, zero-reinforcer situation, i.e. $P(outcome \mid context) = 0$. When they are then exposed to the instrumental contingency, such that $P(outcome \mid action + context) > 0$, this prior experience may enhance their ability to detect the instrumental contingency $\Delta P = P(outcome \mid action) - P(outcome \mid no\ action)$. This interpretation is also supported by the demonstration that delivering "free" rewards (not contingent upon any response of the subject) during the contextual pre-exposure reduces the beneficial effect of this pre-exposure on instrumental learning (Dickinson *et al.*, 1992); by increasing $P(outcome \mid context)$, this reduces the subject's ability to detect the contingency (Wagner *et al.*, 1968; Dickinson & Charnock, 1985; Colwill & Rescorla, 1986). Thus, the formation of context–outcome associations may explain at least some of the ability of action–outcome delays to retard instrumental learning.

3.5.2 Contribution of the hippocampus to instrumental conditioning in the absence of response–reinforcer delays

In the present study, excitotoxic hippocampal lesions impaired instrumental conditioning on an FR-1 schedule in the absence of response–reinforcer delays. This contrasts with the findings of Corbit *et al.* (2000) that electrolytic lesions of the dorsal hippocampus did not affect the acquisition of instrumental responding on a training schedule that progressed from a FI-20 schedule (in which reinforcement is delivered contingent upon the first response that is more than 20 s after the previous reinforcer) up to a RR-20 schedule [in which P(reinforcer│response) = 0.05], consistent with earlier results (Schmaltz & Isaacson, 1967). Rats with excitotoxic NMDA lesions of the dorsal hippocampus also responded at similar, or greater, rates than sham-operated controls in this training regimen (Corbit *et al.*, 2002). Obviously, the discrepancy between these findings and the present results might be due to the differences in the schedules used (FR-1 versus RR-20) or in the lesion (dorsal + ventral hippocampus versus dorsal hippocampus only). However, it is less likely that the impairment seen in the present study was due to a primary motivational difference: although the H-lesioned rats gained less mass than shams, the food consumption tests showed that they ate as many of the pellets used as reinforcers as shams, and as quickly, suggesting that the impairment observed was not due to reduced primary motivation for the food. It is, however, possible that the impairment represents a rate-dependent impairment (i.e. that the H-lesioned rats in the zero-delay condition were responding at their maximum possible rate).

Furthermore, electrolytic lesions of the dorsal hippocampus have been shown to render rats insensitive to changes in the instrumental action–outcome contingency—but in a very specific manner (Corbit & Balleine, 2000; Corbit *et al.*, 2002). One way to test subjects' sensitivity to this contingency is to train them to respond on two levers for two different outcomes, and then to deliver one of the outcomes noncontingently, as well as contingent upon the response. Subjects that are sensitive to the action–outcome contingency should selectively reduce their responding for the foodstuff being delivered noncontingently (Balleine & Dickinson, 1998). Electrolytic dorsal hippocampal lesions impaired this ability, though not the ability to discriminate the two foodstuffs or to respond to changes in their value (Corbit & Balleine, 2000; Corbit *et al.*, 2002). This may have been because the lesion affected contextual conditioning: if an animal cannot associate noncontingent rewards with the context, it may erroneously associate them with its own action. However, excitotoxic lesions of the dorsal hippocampus did not produce this effect, which was reproduced instead by excitotoxic lesions of the entorhinal cortex and subiculum (Corbit *et al.*, 2002); lesions of these structures have also been shown to impair contextual conditioning in Pavlovian tasks (Corodimas & LeDoux, 1995; Maren & Fanselow, 1997; Maren, 1999; Sacchetti *et al.*, 1999; Bucci *et al.*, 2000; Burwell *et al.*, 2004) (though see Phillips & LeDoux, 1995; Bannerman *et al.*, 2001).

3.5.3 Contribution of the hippocampus to instrumental conditioning in the presence of response–reinforcer delays

In contrast, when a delay was imposed between responding and reinforcement in an FR-1 schedule, H-lesioned rats were not impaired at instrumental conditioning, and were even somewhat facilitated in learning, relative to shams, when the reinforcer was delayed by 20 s. Since H-lesioned rats were impaired in the absence of delays, this indicates a delay-dependent improvement in learning, relative to shams. Furthermore, asymptotic rates of responding were reduced less by the delay in H-lesioned rats than in controls. The facilitation of learning after a lesion strongly suggests that the lesion has disrupted one process or strategy that normally competes with another process involved in solving the task (e.g. Jaffard & Meunier, 1993; Kim & Baxter, 2001). Given the involvement of the hippocampus in contextual condi-

tioning (Selden *et al.*, 1991; Kim & Fanselow, 1992; Phillips & LeDoux, 1992; Honey & Good, 1993; Jarrard, 1993; Kim *et al.*, 1993; Phillips & LeDoux, 1995; Maren & Fanselow, 1997; Anagnostaras *et al.*, 1999; Holland & Bouton, 1999; Good, 2002; Rudy *et al.*, 2002), the most obvious explanation is that the lesions facilitated instrumental conditioning with delayed reinforcement by reducing competition from context–reinforcer associations that normally hinder the formation or expression of response–reinforcer associations.

Certain simple explanations of the present results can be ruled out. The delay-dependent impairment makes an explanation in terms of differences in primary motivation for the food *per se* unlikely, and the use of a control (inactive) lever means that differences in responding were not attributable to differences in general activity levels, but instead to the contingencies in force on the active lever. The results also indicated that when there were programmed delays to reinforcement, H-lesioned animals experienced shorter response–reinforcer collection delays, partly because they collected the reinforcer more promptly than shams. This effect probably improved learning in the delay conditions. However, in addition to this effect, there was a further delay-dependent improvement exhibited by H-lesioned rats: even allowing for the shorter response–collection delays that they experienced, their instrumental learning was impaired less by delays than that of sham-operated controls.

The role of the hippocampus in learning an instrumental response with delayed reinforcement has been examined before, though in a very different way. Port *et al.* (1993) found that aspirative lesions of the dorsal hippocampus did not impair appetitive instrumental conditioning with delayed reinforcement. Numerically, lesioned rats were slightly faster to learn (to reach a criterion number of reinforced responses) than shams, but the difference was not significant. This is certainly consistent with the present results, in which a delay-dependent improvement was seen as a result of excitotoxic lesions of the dorsal and ventral hippocampus. However, direct comparison is difficult. Firstly, the lesion extent was different. Secondly, aspirative lesions destroy not just overlying cortex but also fibres of passage (axons of non-hippocampal neurons traversing the hippocampus) (Jarrard & Meldrum, 1993). Thirdly, the task used by Port *et al.* (1993) was quite different to that used in the present study: lever presses led to the delivery of responses after a 5-s delay, while responses during the delay were not reinforced; thus, higher rates of responding inevitably reduced the action–outcome contingency. Fourthly, no other delays were tested and no zero-delay condition was used, so any delay-dependent changes would not have been apparent. Fifthly, no control lever was present, so that responding could only be compared across conditions (inferences from a single condition being potentially confounded with general activity levels). Finally, an autoshaping procedure was used to train the rats, and autoshaping is itself known to be impaired by hippocampal lesions (Reilly & Good, 1989; Good & Honey, 1991; Richmond & Colombo, 2002), so this may have mitigated against finding an improvement in the lesioned group.

3.5.4 Contrasting the effects of hippocampal lesions on instrumental and Pavlovian conditioning involving delayed reinforcement

Hippocampal lesions have also been found to affect Pavlovian conditioning involving temporally non-contiguous stimuli. As discussed above, the two main paradigms used for this purpose are delay conditioning (in which the CS and US overlap and are contiguous) and trace conditioning (in which there is a gap between the CS and the US and they are non-contiguous). Trace conditioning is clearly more analogous to the instrumental conditioning task with delayed reinforcement used in the present study, which had response–reinforcer gaps (Mackintosh, 1974). Hippocampal lesions have been reported to impair trace conditioning to a discrete explicit CS, sparing delay (contiguous) conditioning (Solomon *et al.*,

1986; McEchron *et al.*, 1998; McEchron & Disterhoft, 1999; Weiss *et al.*, 1999a; Beylin *et al.*, 2001; Christian & Thompson, 2003; Shors, 2004). Trace discrimination learning can also be impaired by hippocampal lesions (Schmitt *et al.*, 2004). Thus, hippocampal lesions appear to impair the acquisition of a Pavlovian CR when there is a gap between CS and US. Indeed, it has been suggested that the hippocampus is particularly important for associating discontiguous events—that is, when there is a temporal, or indeed spatial, discontiguity or gap between two events to be associated (Wallenstein *et al.*, 1998). In the current study, however, hippocampal lesions delay-dependently facilitated (relative to shams) the acquisition of an instrumental response when there was a gap between action and outcome. What accounts for these apparently contradictory results?

Firstly, the neural differences between trace and delay conditioning may be not as great as it first seems. The fact that hippocampal lesions have often been shown to impair trace conditioning, but not delay conditioning, may be because trace conditioning is more difficult. If delay conditioning is rendered as hard as trace conditioning by extending the delay ("long-delay conditioning"), the hippocampus is required (Beylin *et al.*, 2001); similarly, H-lesioned subjects can exhibit trace conditioning if pretrained in a delay conditioning paradigm (Beylin *et al.*, 2001; Shors, 2004), some trace discrimination tasks are intact after hippocampal lesions (Savage *et al.*, 2004), and the hippocampus is not required for expression of the trace response after learning (Takehara *et al.*, 2002).

Secondly, it may be that a representation of context can help trace (and perhaps long-delay) Pavlovian conditioning, while contextual associations only hinder instrumental responding in the present task, so a lesion that disrupts contextual processing has a differential effect on the two. For example, if the CS during trace conditioning acts as an occasion setter, signalling that the ensuing context will be followed by a US (an example of feature-positive discrimination), then hippocampal lesions, which can impair both feature-positive discrimination (Holland *et al.*, 1999) and contextual conditioning, might be expected to impair trace conditioning more than delay conditioning (N.J. Mackintosh, personal communication, 12 October 2004). Yet in instrumental conditioning with delayed reinforcement, as examined here, context–outcome associations can only hinder the learning of response–outcome associations, so hippocampal lesions might be expected to improve learning in a delay-dependent fashion, as was observed.

A study by Desmedt *et al.* (2003) supports the hypothesis that trace and delay conditioning endow the context with qualitatively different roles, though it does not allow a unifying conclusion to be drawn about hippocampal function. Mice were trained on a contextual discrimination paradigm: they were given electric shocks in one context (active context: A), but not in another (neutral context: N). The shocks (USs) followed a discrete tone CS (here labelled C), with either a delay conditioning design (C → US) or a trace conditioning design (C → 30 s → US). In the neutral context, CSs were given, but no USs. They were then tested for their CRs to the two contexts (A versus N), and to the CS given in either context (AC versus NC). It would be expected that mice learn more about the context–US association in the trace condition, but more about the CS–US association in the delay condition—and normal mice did indeed learn the context discrimination much faster and better (A > N) when trace conditioning was used; furthermore, trace-conditioned subjects responded more to the context than to the context + CS combination (A > AC), while delay-conditioned subjects did not. Delay-conditioned subjects were better at discriminating the contexts in the presence of the CS than in its absence (AC − NC > A − N). Desmedt *et al.* argue that two processes occur here: in the trace condition, subjects form direct context–US associations, while in the delay condition, the context acts as an occasion setter for the tone (Holland & Bouton, 1999), with context A signalling that C will be followed by shock, and context N signalling that it will not—or perhaps vice versa, with the tone acting as an occasion setter for the context. In their study, excitotoxic hippocam-

pal lesions appeared to *facilitate* occasion setting (improved AC/NC discrimination in the delay group, with AC/NC discrimination now being better than A/N discrimination in the trace group) while impairing direct context processing (impaired A/N discrimination in the trace group). Nevertheless, although contexts clearly have several associative functions, the contribution to the hippocampus to these processes is by no means clear cut (Holland & Bouton, 1999; Holland *et al.*, 1999; Heldt *et al.*, 2002).

3.5.5 Effect of hippocampal lesions on choice involving delayed reinforcement

If lesions of the hippocampus reduce the normal deleterious effects of delays on the ability to associate actions with their outcomes, it might be expected that they would also improve subjects' ability to choose a delayed, large reward in preference to an immediate, small reward. Instead, the opposite pattern of results was observed: postoperatively, H-lesioned rats made impulsive choices, preferring the immediate, small reward. This preference was flexible, responding to changes in the contingencies within the task, and all subjects readily reverted to choosing the large reward when all delays were removed, indicating that they could discriminate the large and small rewards and continued to prefer the large reward when it was not delayed. Upon reintroduction of the delays, however, the preference for the large reward collapsed in the H-lesioned group much more prominently than in shams, indicating that they were less tolerant of delays to reinforcement.

The present results are also somewhat similar to those reported by Rawlins *et al.* (1985), who examined choice between certain and uncertain rewards. Normal rats preferred immediate certain reward to immediate uncertain reward, and also preferred delayed certain reward to immediate uncertain reward; however, rats with hippocampal or medial septal lesions were less tolerant of the delay (or more tolerant of the uncertainty), preferring immediate uncertain reward to delayed certain reward.

Some simple explanations for this effect may readily be ruled out. It is unlikely that lesioned rats' impulsive choice was caused by lower motivation to obtain the food, for two reasons. First, in a separate test, there were no differences in the rate at which they consumed the sucrose pellets used as the reinforcer. Second, the performance of H-lesioned rats was not comparable in other respects to that of a subject with lower primary motivation, such as a sated rat (Cardinal *et al.*, 2000); for example, they did not make more omissions than sham-operated controls. It is also unlikely that H-lesioned rats' preference for the small, immediate reward were the consequence of a positional bias away from the lever producing large, delayed reward: when the delays were omitted, all the H-lesioned rats readily and consistently chose the large reinforcer, only to prefer the small reinforcer again when delays were re-introduced. Furthermore, it is difficult to see that impaired contextual conditioning could explain the pattern of results; as discussed above, the absence of context–reinforcer associations should help, rather than hinder, the ability to associate actions with delayed outcomes. Likewise, although context–response associations may influence instrumental responding and contexts may act as occasion setters (signalling the operation of a particular action–outcome contingency), there is no *a priori* reason to believe that the context should differentially cue instrumental responding more when the outcome is delayed; neither is there conclusive evidence suggesting that hippocampal lesions impair such a process (see Holland & Bouton, 1999).

One obvious difference between the two experiments is that in Experiment 1, lesions were made before training, whereas in Experiment 2, lesions were made after training. Hence it is possible that hippocampal lesions selectively impair the *retrieval* of a well-learned instrumental response or action–outcome contingency inivolving delayed outcomes, while sparing those involving immediate outcomes. However, there are two reasons why this scenario is unlikely. Firstly, animals must be able to perform an action in order for that action to be conditioned instrumentally. Experiment 1 demonstrated that H-lesioned rats

were able to acquire instrumental responses for delayed reinforcement at least as well as, if not better than, sham-operated controls. Therefore, it is unlikely that hippocampal lesions selectively impair the *performance* of instrumental responses for delayed outcomes. Secondly, this idea would not explain why H-lesioned rats showed a reduced preference for the large, delayed reinforcer upon reintroduction of delays, *after* they had shown a strong postoperative preference for the large reinforcer when delays had been consistently omitted. This required new learning on their part, and again Experiment 1 demonstrated that learning with delayed reinforcement is normal in H-lesioned rats, a result that would predict self-controlled rather than impulsive choice upon reintroduction of delays.

The task used does not determine whether H-lesioned rats exhibit altered sensitivity to reinforcer magnitude or delay, for either abnormality might produce impulsive choice (Ho *et al.*, 1999) (see e.g. p. 41). Although H-lesioned rats were able to discriminate in absolute terms between the large and the small reinforcer (consistent with previous studies, e.g. Kesner & Williams, 1995), it is possible that they discriminated between the reinforcer magnitudes to a lesser extent than shams. In this scenario, H-lesioned rats might have exhibited impulsive choice simply because the perceived value of the large reinforcer was not subjectively big enough to compensate for the normal effects of the delay. Alternatively, H-lesioned rats may perceive reward magnitudes normally, and exhibit impulsive choice because they are specifically hypersensitive to (intolerant of) the effects of delays to reinforcement. Such evidence as exists suggests that H-lesioned rats perceive reward magnitude normally (Kesner & Williams, 1995; Gilbert & Kesner, 2002). Experiment 1 indicated that H-lesioned rats are somewhat better than shams at instrumental conditioning with action–outcome delays of >10 s. This suggests that H-lesioned rats associated the action with the delayed outcome normally in Experiment 2, so if they also perceived its magnitude normally, then it is likely that they valued the delayed outcome less.

The present results may also be explained in terms of altered temporal perception. For example, a lesion that increased the speed of an "internal clock" (Gibbon *et al.*, 1997; Buhusi & Meck, 2005) might affect choice prospectively in this task (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 immediate lever), or might affect retrospective choice (i.e. the lesioned subject experiences the delay to the large reinforcer as longer than it actually is, causing it to value the reinforcer less than shams). However, the evidence for the role of the hippocampus in temporal perception is inconclusive: some studies have found that aspirative hippocampal lesions did not affect timing behaviour (Rawlins *et al.*, 1983; Port *et al.*, 1986; Dietrich *et al.*, 1997; Dietrich & Allen, 1998), whereas others have suggested that lesions of the hippocampus or fimbria/fornix speed up an internal clock, or reduce the estimation of time periods when a stimulus being timed is interrupted (Meck *et al.*, 1984; Olton *et al.*, 1987; Meck, 1988; Hata & Okaichi, 1998; Wallenstein *et al.*, 1998).

Finally, the hippocampus is heavily connected to a number of other structures known to play a role in subjects' relative preference between immediate, small and delayed, large rewards. Lesions of the AcbC, BLA, and OFC have all been found to produce impulsive choice (Cardinal *et al.*, 2001; Kheramin *et al.*, 2002; Mobini *et al.*, 2002; Winstanley *et al.*, 2004b). OFC lesions appear to alter the processing of reward magnitude as well as delay, and lesions here have produced both impulsive and self-controlled choice under different circumstances (Kheramin *et al.*, 2002; Mobini *et al.*, 2002; Kheramin *et al.*, 2003; Winstanley *et al.*, 2004b). AcbC lesions appear to have a more selective effect on the processing of delays, impairing both preference for, and learning with, delayed rewards in the absence of effects on reward magnitude processing (Cardinal *et al.*, 2001; Cardinal & Cheung, 2005) (Chapter 4). Although the hippocampal formation projects heavily to most of the Acb, via the subiculum (Brog *et al.*, 1993), and H-

lesioned rats in the present study showed the impulsive choice known to be exhibited by AcbC-lesioned rats, H-lesioned rats showed the opposite effect to AcbC-lesioned rats in the simple instrumental learning task, being delay-dependently improved rather than impaired relative to shams.

3.6 CONCLUSIONS

These experiments have demonstrated that excitotoxic lesions of the hippocampus ameliorate the deleterious effects of response–reinforcer delays on instrumental learning. H-lesioned rats responded slightly less than controls in the absence of delays, but they became better at learning (relative to shams) as the delays increased, in a delay-dependent fashion. This may have been because the lesion hindered the formation of context–outcome associations, promoting response–outcome association instead. In contrast, lesioned rats exhibited impulsive choice, preferring an immediate, small reward to a delayed, larger reward, even though they preferred the large reward when it was not delayed. Thus, lesioned rats were better at learning with delayed reinforcement but worse at choosing it, suggesting that self-controlled choice and learning with delayed reinforcement tax different psychological processes.

Chapter 4: Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats

4.1 ABSTRACT

Background: Animals must frequently make choices between alternative courses of action, seeking to maximize the benefit obtained. They must therefore evaluate the magnitude and the likelihood of the available outcomes. Little is known of the neural basis of this process, or what might predispose individuals to be overly conservative or to take risks excessively (avoiding or preferring uncertainty, respectively). The nucleus accumbens core (AcbC) is known to contribute to rats' ability to choose large, delayed rewards over small, immediate rewards; AcbC lesions cause impulsive choice and an impairment in learning with delayed reinforcement. However, it is not known how the AcbC contributes to choice involving probabilistic reinforcement, such as between a large, uncertain reward and a small, certain reward. The present experiments examined the effects of excitotoxic lesions of the AcbC on probabilistic choice in rats.

Results: Rats chose between a single food pellet delivered with certainty ($p = 1$) and four food pellets delivered with varying degrees of uncertainty ($p = 1, 0.5, 0.25, 0.125$, and 0.0625) in a discrete-trial task, with the large-reinforcer probability decreasing or increasing across the session. Subjects were trained on this task and then received excitotoxic or sham lesions of the AcbC before being retested. After a transient period during which AcbC-lesioned rats exhibited relative indifference between the two alternatives compared to controls, AcbC-lesioned rats came to exhibit risk-averse choice, choosing the large reinforcer less often than controls when it was uncertain, to the extent that they obtained less food as a result. Rats behaved as if indifferent between a single certain pellet and four pellets at $p = 0.32$ (sham-operated subjects) or at $p = 0.70$ (AcbC-lesioned subjects) by the end of testing. When the probabilities did not vary across the session, AcbC-lesioned rats and controls strongly preferred the large reinforcer when it was certain, and strongly preferred the small reinforcer when the large reinforcer was very unlikely ($p = 0.0625$), with no differences between AcbC-lesioned and sham-operated groups.

Conclusions: These results support the view that the AcbC contributes to action selection by promoting the choice of uncertain, as well as delayed, reinforcement.

4.2 BACKGROUND

Animals often need to choose between different courses of action on the basis of the eventual rewarding or reinforcing outcomes of those actions. However, the relationship between an action and an outcome is frequently uncertain: animals do not always obtain that for which they work. Therefore, animals must incorporate information on the probability of obtaining different rewards when making decisions about what to do. Little is known of the neural basis of this process. Furthermore, when making decisions under conditions of uncertainty, individuals vary as to how much uncertainty or risk they are willing to tolerate. Formally, individuals differ in how much they "discount" the value of reinforcers as the uncertainty of the reinforcer increases (i.e. as the probability of the reinforcer declines, or the odds against obtaining the

reinforcer increase) (Ho *et al.*, 1999). Risk taking is one aspect of the personality trait of impulsivity (Daruna & Barnes, 1993; Eysenck, 1993; Evenden, 1999a) and is a feature of a number of psychiatric disorders, including pathological gambling and certain personality disorders (Roy *et al.*, 1989; Coccaro & Siever, 1995; APA, 2000; Holt *et al.*, 2003). The term "risk" implies exposure to the possibility of an aversive consequence (OUP, 1997), which may include the possibility of not obtaining an anticipated reward. In the appetitive domain, risk taking is exemplified by the tendency to choose large rewards that are very uncertain, in preference to smaller, certain rewards. Abnormal risk taking may reflect dysfunction of reinforcement learning systems that mediate the effects of uncertain reward or punishment.

The Acb is one candidate structure that may influence choice involving uncertainty. The Acb responds to anticipated rewards in humans, other primates, and rats (Schultz *et al.*, 1992; Miyazaki *et al.*, 1998; Martin & Ono, 2000; Schultz *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; Cromwell & Schultz, 2003; Bjork *et al.*, 2004), and is innervated by dopamine (DA) neurons that respond to errors in reward prediction in a manner appropriate for a teaching signal (Schultz *et al.*, 1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000). There is clear evidence that the Acb is involved in the processing of delayed reinforcement and its influence upon choice. Damage to the AcbC produces impulsive choice in rats (Cardinal *et al.*, 2001; 2003b), reducing their ability to choose large, delayed rewards in preference to small, immediate rewards, yet these and other similar lesions do not appear to impair rats' ability to discriminate reward size (Balleine & Killcross, 1994; Salamone *et al.*, 1994; Brown & Bowman, 1995; Cousins *et al.*, 1996; Salamone *et al.*, 2001; Cardinal *et al.*, 2003b; 2004; Giertler *et al.*, 2004; Cardinal & Cheung, 2005) (Chapter 4). Furthermore, AcbC lesions impair rats' ability to learn instrumental actions when the outcomes of those actions are delayed (Cardinal & Cheung, 2005) (Chapter 4). The Acb may also be involved in the processing of uncertain or probabilistic reinforcement. DA neurons that innervate the Acb may fire in a manner related to reward probability (Fiorillo *et al.*, 2003; Niv *et al.*, 2005; Tobler *et al.*, 2005) and the midbrain, the site of the cell bodies of these neurons, responds to stimulus uncertainty in humans (Aron *et al.*, 2004). A greater BOLD response is observed in the human Acb during the selection of high-reward/high-risk options, compared to low-reward/low-risk outcomes, in a task where the risk is of not winning (Ernst *et al.*, 2004), with similar activation to high-reward/high-risk option selection in a task where the risk is of losing (Matthews *et al.*, 2004); this latter activation was correlated with personality measures of harm avoidance. However, these studies are correlative, and it is not known whether the AcbC is causally involved in regulating choice involving uncertain reinforcement.

The present study sought to examine the contribution of the AcbC to choice involving probabilistic reinforcement in rats. Rats were trained on a task in which they could choose regularly between a certain, small reward and an uncertain, large reward in discrete trials (Figure 45, p. 110); excitotoxic AcbC lesions were then made before the rats were retested postoperatively. Preoperatively, the proportion of choice trials in which the large reinforcer was chosen was approximately a linear function of the large-reinforcer probability. Postoperatively, after a transient period in which AcbC-lesioned rats were relatively indifferent between the two reinforcers, compared to sham-operated controls, a stable state emerged in which AcbC-lesioned rats chose the large, uncertain reinforcer less often than shams did. This pattern persisted regardless of whether the large-reinforcer probability increased or decreased across the session. AcbC-lesioned rats and controls continued to exhibit a strong preference for the large reinforcer when it was consistently certain, and a strong preference for the small, certain reinforcer when the large reinforcer was very unlikely; the lesioned and sham-operated groups did not differ from each other in either of these conditions. These results suggest that the AcbC is necessary for the normal impact of unlikely (as well as delayed) reinforcers upon choice.
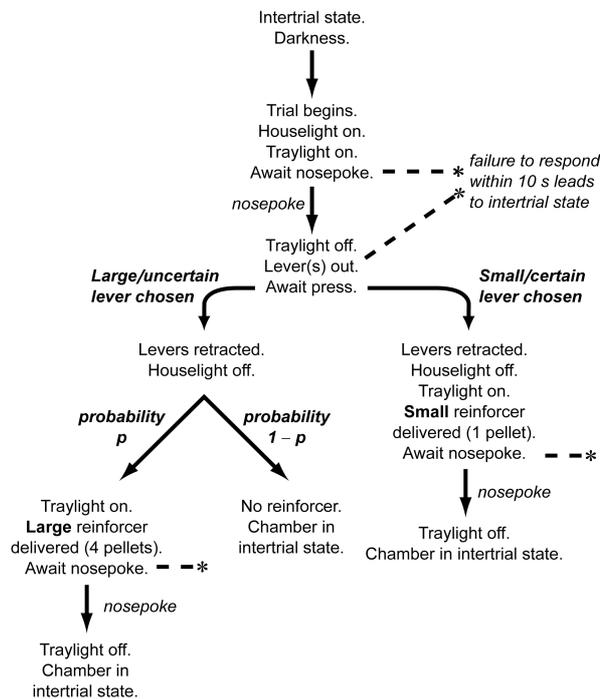
Figure 45: Task schematic: choice between small, certain and large, uncertain rewards

Probabilistic choice task, based on similar tasks involving choice between delayed reinforcers (Evenden & Ryan, 1996; Cardinal *et al.*, 2000). Hungry rats regularly chose between two levers. Responding on one lever led to the certain delivery of a small food reward (1 pellet); responding on the other led to a much larger food reward (4 pellets), but this reward was uncertain, and was delivered with a probability (*p*) ranging from 1 to 0.0625. The figure shows the format of a single trial. Trials began at regular intervals (every 40 s). Sessions consisted of 5 blocks. In each block, 16 single-lever trials were given (8 trials for each lever, randomized in pairs), to ensure the animals sampled the options available at that time; these were followed by 10 choice trials. The probability of delivery of the large reinforcer was varied systematically across the session: probabilities for each block were initially 1, 0.5, 0.25, 0.125, and 0.0625, respectively (see Table 6, p. 113).

4.3 METHODS

4.3.1 Subjects and housing conditions

The subjects were 24 male Lister hooded rats (Harlan-Olac UK Ltd) housed in a temperature-controlled room (minimum 22°C) under a 12:12 h reversed light–dark cycle (lights off 07:30 to 19:30). Subjects were approximately 15 weeks old on arrival at the laboratory and were given a minimum of a week to acclimatize, with free access to food, before experiments began. Preoperatively, subjects were housed in pairs; postoperatively, they were housed individually. Experiments took place between 09:00 and 21:00, with individual subjects being tested at a consistent time of day. Subjects had free access to water. During behavioural testing, subjects were fed ~15–16 g/day, an amount that maintains ~85–90% of free-feeding mass in normal male Lister hooded rats (the free-feeding mass being a steadily increasing quantity at this age). Feeding occurred in the home cages at the end of the experimental day. As it was possible for subjects to earn substantial amounts of food in the behavioural tasks, the amount of food actually earned was taken into account when feeding with the maintenance diet in the home cages. All procedures were subject to UK Home Office approval (Project Licence 80/1767) under the Animals (Scientific Procedures) Act 1986.

4.3.2 Behavioural apparatus

Behavioural testing was conducted in one of two types of operant chamber of identical configuration (from Med Associates Inc., Georgia, Vermont, USA, or Paul Fray Ltd, Cambridge, UK). Each chamber was fitted with a 2.8 W overhead house light and two retractable levers on either side of an alcove fitted with an infrared photodiode to detect head entry and a 2.8 W lightbulb ("traylight"). Sucrose pellets (45 mg, Rodent Diet Formula P, Noyes, Lancaster, New Hampshire, USA) could be delivered into the alcove. The chambers were enclosed within sound-attenuating boxes fitted with fans to provide air circulation. The apparatus was controlled by software written by RNC in C++ (Stroustrup, 1986) using the Whisker control system (Cardinal, 2000; Cardinal & Aitken, 2001). Equal

numbers of subjects were trained in the two brands of operant chamber (12 subjects in each type). Individual subjects were always tested in the same operant chamber.

4.3.3 Initial training

Rats were first trained to press the left lever for single pellets on an FR-1 schedule, in 30-min sessions, until they had obtained a total of 100 pellets. This procedure was repeated for the right lever. They were then trained to nosepoke to initiate presentation of a lever in discrete trials. Each session began with the levers retracted and the operant chamber in darkness. Every 40 s, a trial began with illumination of the houselight and the traylight. The subject was required to make a nosepoke response within 10 s, or the current trial was aborted and the chamber returned to darkness. If the subject nosepoked within this time limit, the traylight was extinguished and a single lever presented. If the rat failed to respond on the lever within 10 s, the lever was retracted and the chamber darkened, but if it responded, the houselight was switched off, a single pellet was delivered immediately and the traylight was illuminated until the rat collected the pellet (or until a 10-s collection time limit elapsed, whereupon the chamber was darkened). In every pair of trials, the left lever was presented once and the right lever once, though the order within the pair of trials was random. The schedule was implemented in the ImpulsiveChoice program (Cardinal, 2002a). Rats were trained to a criterion of 60 successful trials in one hour (the maximum possible with a 40-s period being 90). They then proceeded to the full task.

4.3.4 Probabilistic choice task

The task was based on delayed reinforcement choice tasks that have been described before (Evenden & Ryan, 1996; Cardinal *et al.*, 2000). The session began in darkness with the levers retracted; this was designated the intertrial state. Trials began at 40-s intervals; the format of a single trial is shown in Figure 45 (p. 110). Each trial began with the illumination of the houselight and the traylight. The rat was required to make a nosepoke response, ensuring that it was centrally located at the start of the trial (latency to poke was designated the initiation latency). If the rat did not respond within 10 s of the start of the trial, the operant chamber was reset to the intertrial state until the next trial began and the trial was scored as an omission. If the rat was already nosepoking when the trial began, the next stage followed immediately. Upon a successful nosepoke, the traylight was extinguished and one or both levers were extended. One lever was designated the Large/Uncertain lever, the other the Small/Certain lever (counterbalanced left/right). The latency to choose a lever was recorded. (If the rat did not respond within 10 s of lever presentation, the chamber was reset to the intertrial state until the next trial and the trial was scored as an omission.) When a lever was chosen, both levers were retracted and the houselight was switched off. Choice of the Small lever caused the certain delivery of one pellet; choice of the Large lever caused the delivery of 4 pellets with a particular probability (see below). When reinforcement was delivered, the traylight was switched on. Multiple pellets were delivered 0.5 s apart. If the rat collected the pellets before the next trial began, then the traylight was switched off and the time from delivery of the first pellet until a nosepoke occurred was recorded as the collection latency. If the rat did not collect the food within 10 s of its delivery, the operant chamber entered the intertrial state, though collection latencies were still recorded up to the start of the next trial. The chamber was then in the intertrial state and remained so until the next trial. There was no mechanism to remove uneaten pellets, but failure to collect the reward was an extremely rare event. The large-reinforcer probability was varied systematically across the session as follows. A session consisted of 5 blocks, each comprising 16 trials in which only one lever was presented (8 trials for each lever, randomized in pairs) followed by 10 free-choice trials. The probability that the large reinforcer was delivered, given that the Large lever had been chosen ($p_{reinforcer}$), varied across blocks: it was initially 1, 0.5, 0.25, 0.125, and 0.0625, respectively, for each block. As trials began every 40 s and there were 130 trials per session, the total session length was ~87 minutes; subjects received one session per day. Choice ratios (percentage choice of the large reinforcer, for each trial block) were calculated using only choice trials on which the subject responded. The schedule was implemented in the ImpulsiveChoice program (Cardinal, 2002a).

4.3.5 Excitotoxic lesions of the AcbC

Subjects were anaesthetized with Avertin (2% w/v 2,2,2-tribromoethanol, 1% w/v 2-methylbutan-2-ol, and 8% v/v ethanol in PBS, sterilized by filtration, 10 ml/kg intraperitoneally) and placed in a Kopf or Stoelting stereotaxic frame (David Kopf Instruments, Tujunga, California, USA; Stoelting Co., Wood Dale, Illinois, USA) fitted with atraumatic ear bars. The skull was exposed and a dental drill was used to remove the bone directly above the injection sites. The dura mater was broken with the tip of a hypodermic needle, avoiding damage to underlying venous sinuses. Excitotoxic lesions of the AcbC were made by injecting 0.5 µl of 0.09 M quinolinic acid (Sigma, UK) per side through a glass micropipette at coordinates 1.2 mm anterior to bregma, ±1.8 mm from the midline, and 7.1 mm below the skull surface at bregma; the incisor bar was 3.3 mm below the interaural line (Paxinos & Watson, 1998). The toxin had been dissolved in 0.1 M phosphate buffer (composition 0.07 M Na_2HPO_4, 0.028 M NaH_2PO_4 in double-distilled water, sterilized by filtration) and adjusted with NaOH to a final pH of 7.2–7.4. Toxin was injected over 3 min and the micropipette was left in place for 2 min following injections. Sham lesions were made in the same manner except that vehicle was infused. At the end of the operation, animals were given 15 ml/kg of sterile 5% w/v glucose, 0.9% w/v sodium chloride intraperitoneally. They were given a week to recover, with free access to food, and were handled regularly. Any instances of postoperative constipation were treated with liquid paraffin orally and rectally. At the end of this period, food restriction commenced or was resumed.

4.3.6 Postoperative testing

Subjects were trained preoperatively and tested postoperatively according to the schedule shown in Table 6. In the basic task, used for preoperative training, the probability of large reinforcer delivery declined across trial blocks from 1 to 0.0625 (in the order 1, 0.5, 0.25, 0.125, 0.0625). After subjects had been tested postoperatively for 12 sessions on this schedule, satiety tests were given, to establish the effect of varying primary motivational state on preference for probabilistic reinforcement. Subjects were tested for 4 sessions while alternating between hungry and sated states on consecutive days in counterbalanced fashion (half the subjects experienced hungry and sated days in the order HSHS, and half in the order SHSH). Following a "hungry" session, animals were placed on free food (maintenance diet) until the start of the next day's "sated" session, at which time the food was again removed for the "hungry" session to follow. The comparison was therefore between satiety and food deprivation for ~22 h. Next, subjects were returned to the hungry state and tested for 6 sessions on a schedule in which both the large and small reinforcer were delivered with certainty. Next, the element of uncertainty was reintroduced for another 12 sessions, but this time the probability of large reinforcer delivery (given that the Large lever had been chosen) *increased* across blocks from 0.0625 to 1 (in the order 0.0625, 0.125, 0.25, 0.5, 1). Finally, subjects were tested for 6 sessions with the large reinforcer always being very unlikely ($p = 0.0625$), with the small reinforcer remaining certain.

Sessions	Description	Probability (small reward of 1 pellet)	Probability (large reward of 4 pellets)
1–12	Preoperative training (12 sessions)	$p = 1$	$p = 1$ to 0.0625, decreasing
—	Surgery and recovery	—	—
13–24	Postoperative baseline testing (12 sessions)	$p = 1$	$p = 1$ to 0.0625, decreasing
25–28	Satiety tests (4 sessions)	$p = 1$	$p = 1$ to 0.0625, decreasing
29–34	Equal probabilities, certain (6 sessions)	$p = 1$	$p = 1$
35–46	Increasing probabilities (12 sessions)	$p = 1$	$p = 0.0625$ to 1, increasing
47–52	Large reinforcer always very unlikely (6 sessions)	$p = 1$	$p = 0.0625$

Table 6: Testing schedule for probabilistic choice task

Subjects were trained and tested according to the schedule shown here. Initial pre- and postoperative testing was conducted with the probability of large reinforcer delivery declining across trial blocks from 1 to 0.0625 (the steps were p = 1, 0.5, 0.25, 0.125, and 0.0625). Subsequently, subjects were tested alternating between the hungry and sated state (as described in the Methods), before the reinforcement probabilities were manipulated further, as indicated.

4.3.7 Locomotor activity in a novel environment

Locomotor activity was measured in wire mesh cages, 25 (W) × 40 (D) × 18 (H) cm, each equipped with a water bottle and two horizontal photocell beams situated 1 cm from the floor that enabled movements along the long axis of the cage to be registered. The apparatus was controlled by software written by RNC in Arachnid (Paul Fray Ltd, Cambridge), a real-time extension to BBC BASIC V running on an Acorn Archimedes series computer. Subjects were placed in these cages, which were initially unfamiliar to them, and their activity was recorded for 2 h. All animals were tested in the food-deprived state. Locomotor hyperactivity and reduced body mass gain have previously been part of the phenotype of AcbC-lesioned rats, though without alterations in the consumption of the reinforcer used in the present experiments (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal, 2001; Cardinal *et al.*, 2001; Cardinal & Cheung, 2005) (Chapter 4).

4.3.8 Histology

Rats were deeply anaesthetized with pentobarbitone sodium (200 mg/ml, minimum of 1.5 ml i.p.) and perfused transcardially with 0.01 M PBS followed by 4% paraformaldehyde in PBS. Their brains were removed and postfixed in paraformaldehyde before being dehydrated in 20% sucrose for cryoprotection. The brains were sectioned coronally at 60 μm thickness on a freezing microtome and every third section mounted on chromium potassium sulphate/gelatin-coated glass microscope slides and allowed to dry. Sections were passed through a series of ethanol solutions of descending concentration (3 minutes in each of 100%, 95%, and 70% v/v ethanol in water) and stained for ~5 min with cresyl violet. The stain comprises 0.05% w/v aqueous cresyl violet (Raymond A. Lamb Ltd, Eastbourne, UK), 2 mM acetic acid, and 5 mM formic acid in water. Following staining, sections were rinsed in water and 70% ethanol before being differentiated in 95% ethanol. Finally, they were dehydrated and delipidated in 100% ethanol and Histoclear (National Diagnostics, UK) before being cover-slipped using DePeX mounting medium (BDH, UK) and allowed to dry. The sections were used to verify lesion placement and assess the extent of lesion-induced neuronal loss. Lesions were detectable as the absence of visible neurons (cell bodies of the order of 100 μm in diameter with a characteristic shape and appearance), often associated with a degree of tissue collapse (sometimes with consequent ventricular expansion when the lesion was adjacent to a ventricle) and gliosis (visible as the presence of smaller, densely staining cells).

4.3.9 Data analysis

Data collected by the chamber control programs were imported into a relational database (Microsoft Access 97) for case selection and analysed with SPSS 11. Figures were created with SigmaPlot 2001/v7 and Adobe Illustrator 8.

All graphs show group means and error bars are ±1 standard error of the mean (SEM) unless otherwise stated. Count data (e.g. locomotor activity counts), for which variance increases with the mean, were subjected to a square-root transformation prior to any analysis (Howell, 1997). Homogeneity of variance was verified using Levene's test (Levene, 1960). General linear models are described as *dependent variable = $A_2 \times B_{cov} \times (C_5 \times D_{cov} \times S)$* where A is a between-subjects factor with two levels, B is a between-subjects covariate, C is a within-subjects factor with five levels, and D is a within-subjects covariate; S denotes subjects in designs involving within-subjects factors (Keppel, 1982). For repeated measures analyses, Mauchly's test of sphericity of the covariance matrix was applied (Mauchly, 1940) and the degrees of freedom corrected to more conservative values by multiplying them by the Huynh–Feldt epsilon $\tilde{\varepsilon}$ for any terms involving factors in which the sphericity assumption was violated (Huynh & Feldt, 1970). Where multiple comparisons were conducted *post hoc* following a significant overall ANOVA effect for a factor with more than three levels, *p* values were corrected using the Šidák correction (Šidák, 1967), in which $p_{\text{corrected}} = 1 - (1 - p_{\text{uncorrected}})^n$ for *n* comparisons.

4.4 RESULTS

4.4.1 Histology

There were four postoperative deaths. Histological analysis revealed that the lesions were incomplete or encroached significantly on neighbouring structures in two subjects. These subjects were excluded; final group numbers were therefore 6 (AcbC)[22] and 12 (sham)[23]. Lesions of the AcbC encompassed most of the core subregion; neuronal loss and associated gliosis extended in an anteroposterior direction from approximately 2.7 mm to 0.2 mm anterior to bregma, and did not extend ventrally or caudally into the ventral pallidum or olfactory tubercle. Damage to the ventromedial caudate–putamen was occasionally seen; damage to the AcbSh was restricted to the lateral edge of the dorsal shell. Schematics of the lesions are
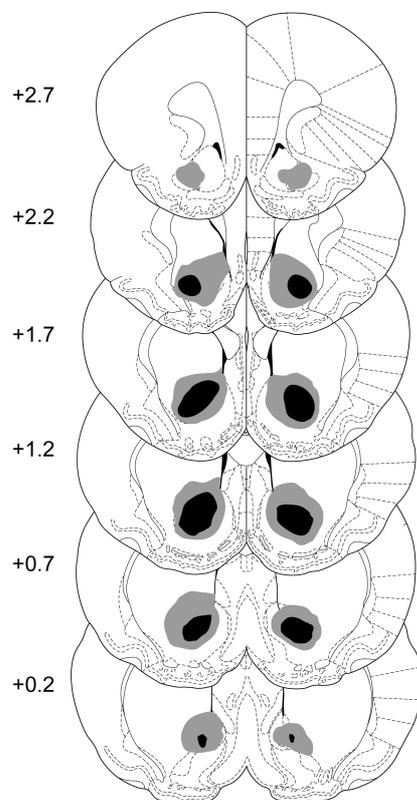
Figure 46: Schematic of lesions of the AcbC

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 (from top to bottom) +2.7, +2.2, +1.7, +1.2, +0.7, and +0.2 mm relative (anterior) to bregma. Diagrams are modified from Paxinos & Watson (1998).

---

[22] AcbC group: final subjects R1, R5, R9, R10, R18, R24 (*n* = 6).
[23] Sham group: final subjects R3, R4, R6, R11, R12, R15, R16, R17, R19, R20, R22, R23 (*n* = 12).

shown in Figure 46. Photomicrographs of lesions with identical parameters have been presented before (Parkinson *et al.*, 1999a; Cardinal, 2001; Cardinal & Cheung, 2005) (see Chapter 4, Figure 26, p. 61).

4.4.2 Preoperative choice

The groups remained matched for preoperative choice behaviour following later histological selection (Figure 47a). Choice ratios (percentage choice of the large reinforcer, for each trial block) calculated across sessions 10–12 (see Table 6, p. 113) were analysed using the model lesion intent$_2$ × (large-reinforcer probability$_5$ × S). There was a robust effect of probability ($F_{3.3,52.9}$ = 70.6, $\tilde{\varepsilon}$ = .826, $p < .001$) but no effect of lesion intent and no lesion intent × probability interaction ($Fs < 1$, NS).
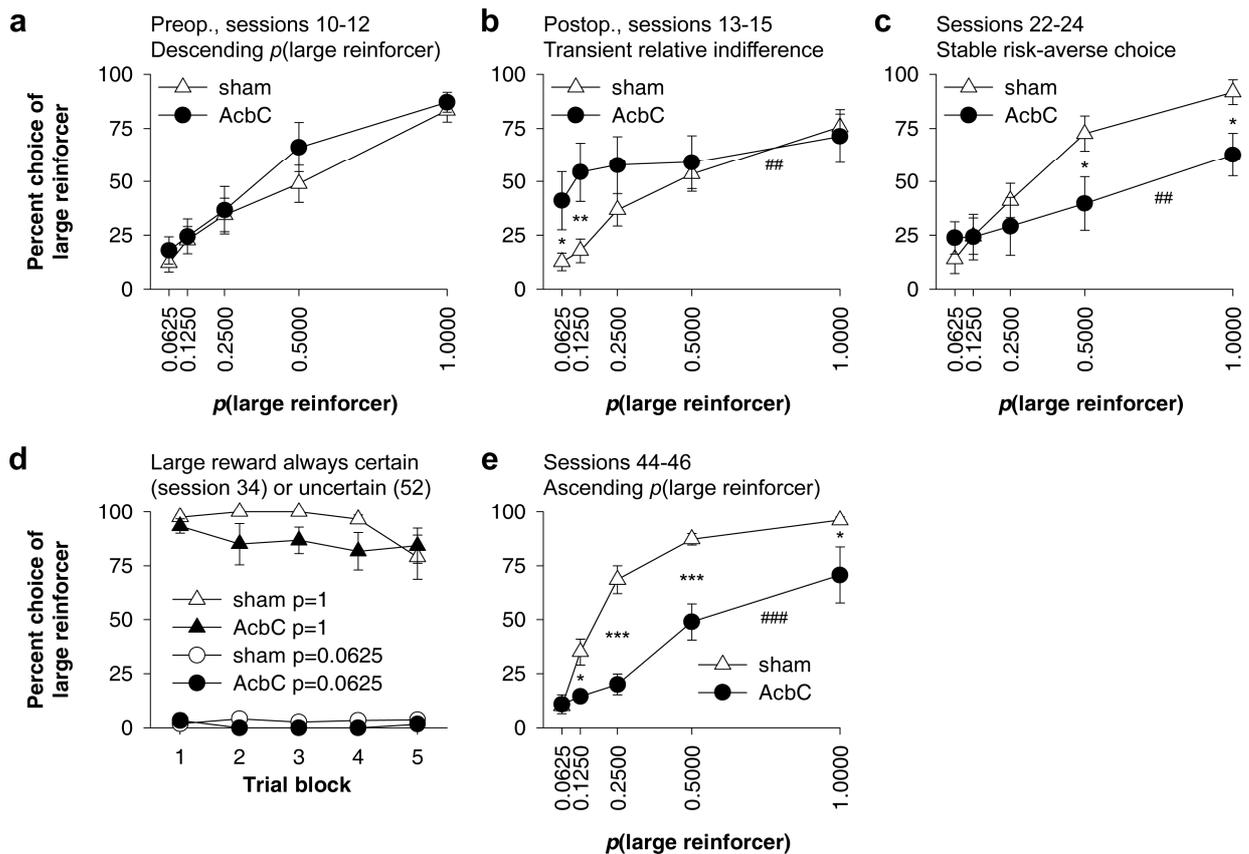
Figure 47: Choice with probabilistic reinforcement in AcbC-lesioned and sham-operated rats

(a) Preoperative patterns of choice. There were no differences between the groups preoperatively. **(b)** The first three postoperative sessions. Transiently, AcbC-lesioned rats exhibited relative indifference between the two alternatives; their preference did not differ significantly from 50% at any large-reinforcer probability. As a result, AcbC-lesioned rats preferred the large, unlikely reinforcer more than shams did when its probability was 0.0625 and 0.125 (## $p <$ .01, lesion × probability interaction; * $p < .05$, ** $p < .01$, comparison to shams at individual probabilities). However, both groups were influenced by the large-reinforcer probability ($p \leq .004$). **(c)** The last three postoperative sessions on the same basic task. By this point, AcbC-lesioned rats preferred the large reinforcer less when its probability was 0.5 or 1 (## $p < .01$, interaction; * $p < .05$, simple effects). Again, both groups were influenced by the large-reinforcer probability ($p < .001$). **(d)** When the four-pellet reinforcer and the one-pellet reinforcer were both certain, all groups preferred the four-pellet reinforcer, and when the four-pellet reinforcer was always very unlikely (delivered with a probability of 0.0625) and the one-pellet reinforcer was certain, all groups preferred the one-pellet reinforcer, with no differences between AcbC-lesioned and sham-operated rats. This indicates that both groups discriminated the reinforcers themselves and discriminated their probability of delivery. **(e)** Choice following further training in which the large-reinforcer probability increased, rather than decreased, across each session. The pattern of choice is similar to **c,** in that AcbC-lesioned rats were risk-averse compared to shams, i.e. less likely to choose the large, unlikely reinforcer (### $p < .001$, interaction; * $p < .05$ and *** $p < .001$, simple effects). The similarity to **c,** despite the reversed task order, also indicates that subjects' choice reflected the probabilities in force rather than the order within a session.

4.4.3 Early postoperative choice

In the initial postoperative period, AcbC-lesioned rats exhibited relative indifference between the two alternatives, choosing the large reinforcer close to 50% of the time at all large-reinforcer probabilities; as a result, AcbC-lesioned rats were more likely than shams to choose the large reinforcer when it was most uncertain (Figure 47b). An analysis of choice ratios calculated across sessions 13–15 was performed using the ANOVA model lesion$_2$ × (probability$_5$ × S). This revealed a lesion × probability interaction ($F_{3.3,53.5}$ = 5.22, $\tilde{\varepsilon}$ = .836, p = .002). Comparison of the two groups at individual large-reinforcer probabilities demonstrated that AcbC-lesioned rats chose the large/uncertain reinforcer more than shams at $p_{\text{reinforcer}}$ = 0.0625 ($p_{\text{statistical}}$ = .02), and at $p_{\text{reinforcer}}$ = 0.125 ($p_{\text{statistical}}$ = .009), but did not differ from shams at reinforcer probabilities of 0.25–1 ($p_{\text{statistical}} \geq$ .158). Nevertheless, simple effects of probability persisted both in shams ($F_{2.8,30.9}$ = 32.3, $\tilde{\varepsilon}$ = .702, $p <$ .001) and in AcbC-lesioned rats ($F_{4,20}$ = 5.37, p = .004). Choice at each $p_{\text{reinforcer}}$ was compared to 50% (indifference) using *post hoc* two-tailed one-sample t tests, correcting $p_{\text{statistical}}$ values using the Šidák correction for 5 comparisons. For shams, choice differed significantly from 50% at large-reinforcer probabilities of 0.0625 (when choice of the large reinforcer was less than 50%), 0.125 (less than 50%), and 1 (greater than 50%) (corrected $p_{\text{statistical}} \leq$ 0.007), but for AcbC-lesioned rats, choice did not differ significantly from 50% at any large-reinforcer probability (corrected $p_{\text{statistical}} \geq$ 0.81).

4.4.4 Final postoperative choice

By the final three sessions of the basic task (sessions 22–24; see Table 6, p. 113), the pattern of choice in AcbC-lesioned rats had changed (Figure 47c). Once more, an analysis of choice ratios using the model lesion$_2$ × (probability$_5$ × S) revealed a lesion × probability interaction ($F_{2.9,46.4}$ = 5.78, $\tilde{\varepsilon}$ = .726, p = .002). By now, however, AcbC-lesioned rats did not differ from shams with reinforcer probabilities of 0.0625–0.25 ($p_{\text{statistical}} \geq$ .386) but chose the large reinforcer *less* than shams when its probability was 0.5 ($p_{\text{statistical}}$ = .037) and 1 ($p_{\text{statistical}}$ = .015). As before, effects of probability persisted both in shams ($F_{2.3,24.9}$ = 49.5, $\tilde{\varepsilon}$ = .565, $p <$ .001) and in AcbC-lesioned rats ($F_{4,20}$ = 9.45, $p <$ .001).

4.4.5 Choice when both reinforcers were certain, or both uncertain

When the large and small reinforcers were both delivered with certainty, AcbC-lesioned and sham-operated rats strongly preferred the large reinforcer; when the small reinforcer was certain and the large reinforcer was consistently unlikely ($p_{\text{reinforcer}}$ = 0.0625), all rats strongly preferred the small reinforcer (Figure 47d). There were no group differences in either case. This indicates that both AcbC-lesioned and sham-operated rats successfully discriminated the large reinforcer from the small reinforcer, and discriminated the certain large reinforcer from the uncertain large reinforcer. Choice ratios from the final sessions of training in these two conditions (sessions 34 and 52; see Table 6, p. 113) were analysed using the model lesion$_2$ × (trial block$_5$ × S). In the "certain" condition (session 34), there was no effect of lesion ($F_{1,15}$ = 2.54, p = .132), no lesion × block interaction (F = 1.42, NS), and no effect of trial block ($F_{1.5,21.9}$ = 2.12, $\tilde{\varepsilon}$ = .365, p = .154). Similarly, in the "uncertain" condition (session 52), there was no effect of lesion (F = 1.35, NS), no lesion × block interaction (F = 1.31, NS), and no effect of trial block ($F <$ 1, NS).

4.4.6 Choice with ascending probabilities

After rats had been trained with the large-reinforcer probability *increasing* across the session, choice behaviour was similar to that with the decreasing-probability version of the task used initially, with AcbC-lesioned rats choosing the large/uncertain reinforcer less often than shams (Figure 47e; compare Figure

47c). Choice ratios from sessions 44–46 (see Table 6, p. 113) were analysed using the model $lesion_2 \times$ $(probability_5 \times S)$. As before, there was a lesion $\times$ probability interaction ($F_{4,64} = 9.29$, $p < .001$), in addition to main effects of lesion ($F_{1,16} = 19.5$, $p < .001$) and probability ($F_{4,64} = 95.6$, $p < .001$), and there were strong effects of probability for both AcbC-lesioned rats ($F_{1,20} = 20.7$, $p < .001$) and shams ($F_{3.1,34.4} = 119.6$, $\tilde{\varepsilon} = .781$, $p < .001$). AcbC-lesioned rats differed from shams at reinforcer probabilities of 0.125 ($p_{\text{statistical}} = .033$), 0.25 ($p_{\text{statistical}} < .001$), 0.5 ($p_{\text{statistical}} < .001$), and 1 ($p_{\text{statistical}} = .013$), but not at $p_{\text{reinforcer}} =$ 0.0625 ($p_{\text{statistical}} = .881$).

4.4.7 Postoperative choice: analysis by experienced probability

Since the task was genuinely probabilistic, and not pseudorandom, it is possible that the probabilities experienced by subjects differed from the programmed probabilities (although experienced probabilities inevitably tend towards programmed probabilities as the number of trials increases). For example, one subject choosing an uncertain reinforcer at $p_{\text{reinforcer}} = 0.5$ for 10 trials might experience 3 rewarded and 7 unrewarded trials (an experienced probability of 0.3), while another might experience 6 rewarded and 4 unrewarded (experienced $p_{\text{reinforcer}} = 0.6$). To establish whether such effects accounted to any degree for the pattern of choice observed in AcbC-lesioned and sham-operated rats, choice was re-analysed for four sets of sessions (preoperative sessions 10–12, early postoperative baseline sessions 13–15, late postoperative baseline sessions 22–24, and sessions 44–46 at the end of training on the increasing-probability version of the task; see Figure 48a–d, and compare to the corresponding programmed-probability versions in Figure 47a–c,e). In each case, choice ratios were analysed using the model $lesion_2 \times$ (experienced $probability_{\text{cov}} \times S$), with the factor $\times$ covariate term included in the model. Experienced probabilities were calculated for all trial types (forced and choice trials), across the sessions concerned.

These analyses confirmed the pattern of results obtained on the basis of programmed probabilities. For the preoperative sessions, as expected, there was a main effect of experienced probability ($F_{1,54} = 319.1$, $p < .001$) but no significant terms involving lesion intent ($Fs < 1$, NS). For the baseline (decreasing-probability) task, both early (sessions 13–15) and late (sessions 22–24) in the postoperative testing, there was a lesion $\times$ experienced probability interaction (early: $F_{1,54} = 25.7$, $p < .001$; late: $F_{1,54} = 20.8$, $p < .001$). For the increasing-probability schedule (sessions 44–46), there was no lesion $\times$ experienced probability interaction ($F_{1,54} = 1.80$, $p = .185$) but there was a main effect of lesion ($F_{1,16.0} = 9.36$, $p = .007$).
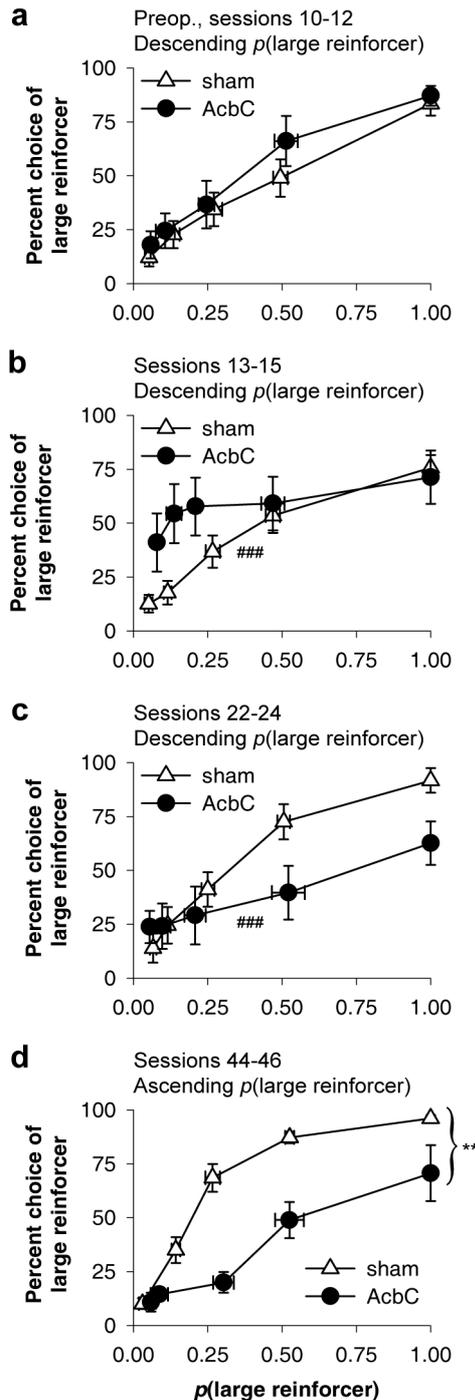
Figure 48: Choice, by experienced probability, in AcbC-lesioned and sham-operated rats

Choice, replotted by experienced (as opposed to programmed) large-reinforcer probabilities. Panels **a–d** correspond to panels **a–c/e** of the previous figure. The statistical patterns of choice remained the same (### $p < .001$, lesion × experienced delay interaction; ** $p < .01$, main effect of lesion).

4.4.8 Indifference probabilities

Choice ratios from sham-operated rats on sessions 22–24 (the final 3 postoperative sessions on the basic task; see Table 6, p. 113) were analysed using four different linear predictors, based either on the probability of delivery of a large reinforcer (given choice of the Large lever), $p_{\text{reinforcer}}$, or of the odds against delivery of a large reinforcer, calculated as *odds against* $= (1 - p_{\text{reinforcer}})/p_{\text{reinforcer}}$. This established that choice patterns were predicted best, in linear fashion, by experienced probabilities (within-subject predictor allowing different slopes for each subject, $r^2 = 0.85$) and programmed probabilities ($r^2 = 0.84$), rather than by experienced odds ($r^2 = 0.61$) or programmed odds ($r^2 = 0.67$). Additionally, optimal behaviour would give choice that was a step function of probability: it is optimal to choose the

small/certain lever whenever the four-pellet reinforcer is delivered with $p_{\text{reinforcer}} < 0.25$ and to choose the large/uncertain lever whenever $p_{\text{reinforcer}} > 0.25$. Therefore, a single-parameter continuous function approximating a step function was also used to predict subjects' choice: the logistic function $y = 100/e^{-(x-m)\div b}$ with y as the percentage choice of the large reinforcer, x as the programmed probability, $b = 0.01$ as an approximation to $b = 0$ and m as the free parameter. However, this gave a poor fit (r^2 calculated as $SS_{\text{model}}/SS_{\text{total}}$ for a nonlinear fit: mean $r^2 = 0.26$; note that individual values of r^2 can fall outside the range [0,1] when calculated this way for nonlinear models) (Cameron & Windmeijer, 1997). Consequently, since choice was best described as a linear function of probability, indifference probabilities were calculated for sham-operated and AcbC-lesioned rats, namely the probability at which rats were equally likely to choose the small/certain and large/uncertain reinforcers. These were calculated via a linear regression of probability on choice (i.e. a regression in which probability was predicted from choice). This method has the potential to produce nonsensical probabilities for individual rats (if, for example, an individual's curve does not go both above and below the 50% choice point in a given set of sessions) but is nonetheless useful for group comparison. Experienced large-reinforcer probabilities (across all types of trials) were used, rather than programmed probabilities, though the pattern of results presented below was not altered by the use of programmed probabilities instead.

The main finding was that by the end of testing, AcbC-lesioned rats had higher indifference probabilities (0.70) than sham-operated rats (0.32) (Figure 49)—that is, while sham-operated rats behaved as if indifferent between a one-pellet certain reinforcer and a four-pellet reinforcer delivered with probability 0.32 (mathematically, an expected number of pellets of $0.32 \times 4 = 1.28$), AcbC-lesioned rats behaved as if indifferent between a one-pellet certain reinforcer and a four-pellet reinforcer delivered with probability 0.70 (an expected number of pellets of 2.8). That is, AcbC-lesioned rats appeared to exhibit risk aversion by the end of testing. The full analysis was as follows. Preoperatively (sessions 10–12), indifference probabilities were 0.43 ± 0.08 (AcbC) and 0.54 ± 0.09 (sham); these did not differ ($F < 1$, NS). In the initial postoperative period (sessions 13–15), indifference probabilities were numerically lower in the lesioned group, being 0.25 ± 0.28 (AcbC) and 0.59 ± 0.12 (sham), but indifference probabilities were highly variable in both groups and these did not differ ($F_{1,16} = 1.76$, $p = .204$). In the later postoperative period (sessions 22–24), indifference probabilities were higher in the lesioned group, being 0.75 ± 0.22 (AcbC) and 0.39 ± 0.15 (sham), though again these did not differ significantly ($F_{1,16} = 1.90$, $p = .187$). In the increasing-probability version of the task (sessions 44–46), indifference probabilities were again higher in the lesioned group, being 0.70 ± 0.15 (AcbC) and 0.32 ± 0.02 (sham). By this stage the difference was highly significant ($F_{1,16} = 12.6$, $p_{\text{statistical}} = .003$), even if corrected for four comparisons ($p_{\text{statistical}}$
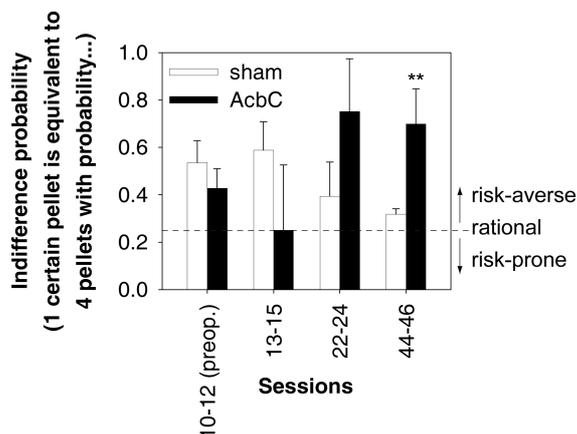
Figure 49: Indifference probabilities in AcbC-lesioned and sham-operated rats

Subjects' behaviour was analysed using a linear regression technique (see text for method of calculation) to estimate the large-reinforcer probability at which they were indifferent between a four-pellet uncertain large reinforcer and a one-pellet certain small reinforcer. Rational choice, and optimal choice in this task, would be an indifference probability of 0.25 (that is, it is rational to be indifferent between a certain one-pellet reinforcer and a four-pellet reinforcer delivered with a probability of 0.25), shown by the dotted line. Lower indifference probabilities imply risk-prone behaviour; higher indifference probabilities imply risk-averse behaviour. Preoperative and successive postoperative indifference probabilities are shown for AcbC-lesioned and sham-operated control rats (** $p < .01$, difference from controls).

= .012) using the Šidák correction.

4.4.9 Omissions and latencies

Omissions were infrequent and not influenced by reinforcer probability or the lesion. Omission data from the final postoperative baseline sessions (sessions 22–24) were analysed. Overall, omissions (either failures to initiate a trial or to respond to an initiated trial) across all trial types occurred at a rate of 2.9 ± 0.9 % (sham) and 5.5 ± 1.9 % (AcbC). Omissions on choice trials for the same sessions were analysed using the model lesion$_2$ × (probability$_5$ × S). There were no effects of lesion ($F_{1,16} = 1.95$, NS) or probability ($F_{1.6,25.2} = 2.56$, $\tilde{\varepsilon} = .394$, $p = .107$), and no interaction ($F = 1.04$, NS). Almost all omissions were failures to initiate a trial (shams 0.9% of choice trials, AcbC 4.4%) rather than failures to respond once a trial had been initiated (shams 0.06% of choice trials, AcbC 0%).

Initiation latencies on choice trials for sessions 22–24 were analysed in the same manner. They were not affected by the lesion ($F < 1$, NS), nor by the large-reinforcer probability ($F_{4,64} = 1.41$, NS), and there was no lesion × probability interaction ($F < 1$, NS).

Response latencies were not affected by the lesion, but were affected both by the time in the session, with responding tending to get slower as the session progressed, and by the likelihood of obtaining a large reinforcer, with responding tending to get faster as large-reinforcer delivery became more likely. Response latencies on choice trials for sessions 22–24 were analysed using the model lesion$_2$ × (trial block$_5$ × choice$_2$ × S). Response latencies varied across trial blocks: response latencies were initially 0.82 s (in the first trial block, when the large-reinforcer probability was 1) and slowed to 1.1 s (in the last trial block, when the large-reinforcer probability was 0.0625) ($F_{3.1,25.1} = 2.97$, $\tilde{\varepsilon} = .785$, $p = .049$). Latencies were not affected by the lesion, or the lever being chosen, and there were no interactions (maximum F was for response: $F_{1,8} = 2.96$, $p = .124$). To establish whether these effects were due to the large-reinforcer probability, or to progressive satiation or the passage of time, data from sessions 44–46 were also analysed, because in these sessions the large-reinforcer probability *increased* within the session. This time, there was a response × trial block interaction ($F_{4,28} = 6.44$, $p = .001$), with no other terms significant (Fs < 1, NS). Responding on the small/certain lever initially took 0.71 s in the first trial block and slowed to 0.95 s in the last trial block ($F_{2.3,24.8} = 3.58$, $\tilde{\varepsilon} = .564$, $p = .038$), but responding on the large/uncertain lever initially took 0.97 s (in the first trial block, when the large-reinforcer probability was 0.0625) and speeded up to 0.79 s (in the last trial block, when the large-reinforcer probability was 1) ($F_{3.5,38.8} = 3.222$, $\tilde{\varepsilon} = .883$, $p = .027$).

The lesion did not affect the latency to collect reward. Food collection latencies on rewarded trials were analysed across sessions 22–24, this time including both forced and choice trials to enable an analysis by response and probability. The model lesion$_2$ × (probability$_5$ × response$_2$ × S) was used; this revealed main effects of response ($F_{1,13} = 13.8$, $p = .003$) and probability ($F_{2.5,32.8} = 3.53$, $\tilde{\varepsilon} = .631$, $p = .031$), but no other significant terms (maximum F was for lesion × response, $F_{1,13} = 3.94$, $p = .069$). Collection was faster following delivery of the large reinforcer than the small (4.1 versus 5.3 s, respectively), and got slightly slower across the session (4.4 s in the first trial block and 4.9 s in the last).

4.4.10 Amount of food obtained

AcbC-lesioned rats obtained less food as a result of their choices (Figure 50a,b). An analysis of the average number of pellets obtained on choice trials in sessions 13–15 using the model lesion$_2$ × (probability$_5$ × S) revealed a main effect of lesion ($F_{1,16} = 8.69$, $p = .009$), as well as an effect of probability ($F_{1.9,30.1} = 97.5$, $\tilde{\varepsilon} = .470$, $p < .001$), but no interaction ($F < 1$, NS). A similar analysis of the final baseline postop-

erative sessions 22–24 revealed a lesion × probability interaction ($F_{2.1,33.8} = 3.29$, $\tilde{\varepsilon} = .529$, $p = .047$) in addition to main effects of lesion ($F_{1,16} = 14.2$, $p = .002$) and probability ($F_{2.1,33.8} = 122.2$, $\tilde{\varepsilon} = .529$, $p < .001$). However, the only probability at which groups significantly differed was $p = 1$ (statistical $p = .014$); when the large reinforcer probability was 0.0625–0.5, the two groups did not differ in the amount of food obtained ($p_{\text{statistical}} \geq .129$).
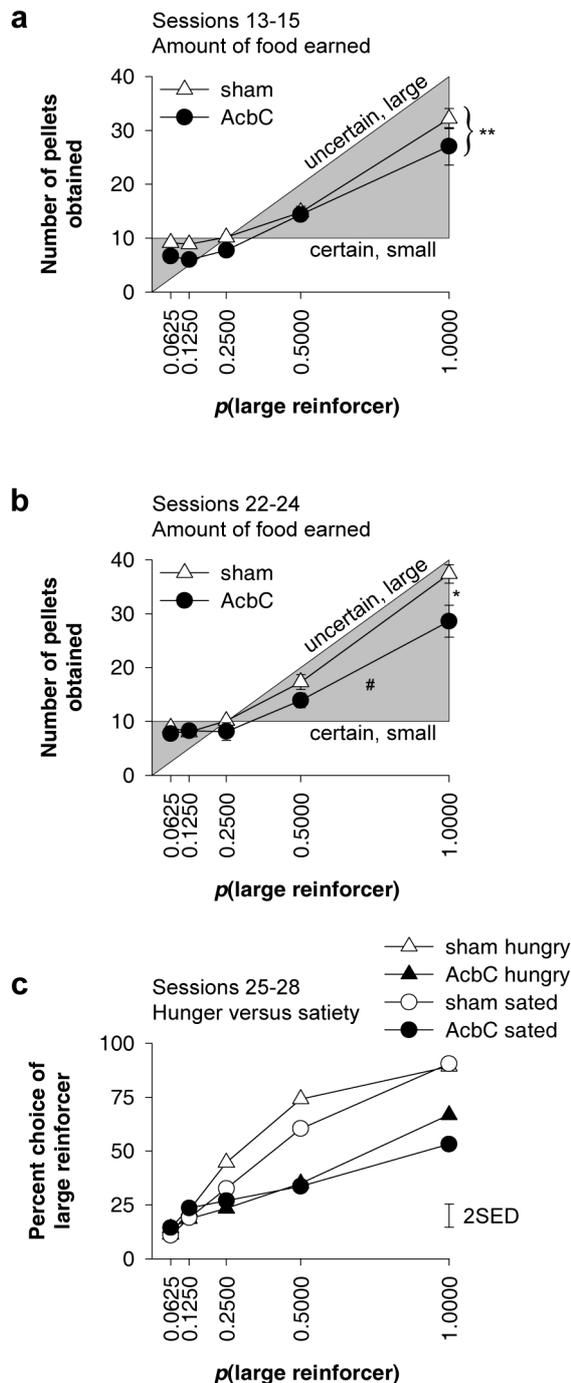
Figure 50: Amount of food obtained, and effects of satiety on choice, in AcbC-lesioned and sham-operated rats

(a) Number of pellets obtained in each trial block; average of the first three postoperative sessions, 13–15 (** $p < .01$, main effect of lesion). The grey area indicates the expected range of options available to a rat making no omissions: consistent responding on the lever delivering the small, certain reward of a single pellet yields 10 pellets per trial block (horizontal border of the grey area); consistent responding on the lever delivering the large, uncertain reward yields an expected number of pellets that varies with the probability in force (as shown by the diagonal border of the grey area). Optimal behaviour, to maximize the expected amount of food, is to choose the small/certain lever when the large (four-pellet) reinforcer probability is less than 0.25 and to choose the large/uncertain lever when this probability exceeds 0.25. **(b)** As for (a), but showing data from the final baseline postoperative sessions, 22–24 (# $p < .05$, lesion × probability interaction; * $p < .05$, simple effect of lesion). **(c)** Effects on choice of alternating subjects between states of hunger and satiety. The error bar is twice the SED for the three-way (lesion × hunger × probability) interaction.

4.4.11 Effects of hunger and satiety on choice

Alternating between hunger and satiety had no substantial effects on choice (Figure 50c). Choice ratios for sessions 25–28 were analysed using the model lesion$_2$ × (hunger$_2$ × probability$_5$ × S). As before, a

main effect of probability ($F_{2.0,32.6} = 38.4$, $\tilde{\varepsilon} = .510$, $p < .001$) and a lesion × probability interaction ($F_{2.0,32.6} = 4.29$, $\tilde{\varepsilon} = .510$, $p = .022$) were present; in addition, there was a marginally significant lesion × hunger × probability interaction ($F_{4,64} = 2.51$, $p = .05$). However, an effect of hunger was not detectable in either group alone, either for shams (hunger: $F_{1,11} = 2.45$, NS; hunger × probability: $F_{4,44} = 2.18$, NS) or for AcbC-lesioned rats (hunger: $F < 1$, NS; hunger × probability: $F_{4,20} = 1.79$, NS). Similarly, the differences between groups persisted both in the hungry (lesion × probability: $F_{2.6,41.7} = 3.66$, $\tilde{\varepsilon} = .652$, $p = .024$) and the sated (lesion × probability: $F_{2.5,39.3} = 4.24$, $\tilde{\varepsilon} = .615$, $p = .016$) conditions.

4.4.12 Locomotor activity and body mass

AcbC-lesioned rats were hyperactive and slower to habituate to a novel environment (Figure 51). AcbC-lesioned rats also gained less mass postoperatively. At the time of surgery, the groups did not differ in mass (shams, 357 ± 4 g; AcbC, 362 ± 6 g; $F < 1$, NS), but at the end of the experiment AcbC-lesioned rats weighed less than shams (shams, 421 ± 7 g; AcbC, 358 ± 10 g; lesion × time, $F_{1,16} = 80.1$, $p < .001$; simple effect of lesion at final time point: $F_{1,16} = 24.5$, $p < .001$). Both effects are consistent with previous results: AcbC-lesioned rats are known to exhibit locomotor hyperactivity (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001; Cardinal & Cheung, 2005) (Chapter 4) and to weigh less than sham-operated controls (Maldonado-Irizarry & Kelley, 1995; Parkinson, 1998; Cardinal *et al.*, 2001; Cardinal & Cheung, 2005) (Chapter 4). They also eat the food used as the maintenance diet in the present study more slowly than sham-operated controls, and eat less of it in a given time, but do not differ in consumption of the sucrose pellets used as reinforcers in the present study (Cardinal, 2001; Cardinal *et al.*, 2001). It is not known whether there are metabolic differences in AcbC-lesioned rats above and beyond the tendency to eat somewhat less and to be hyperactive (though see Kelley, 2004). However, differences in mass between AcbC-lesioned and sham-operated rats are also apparent when they have been fed *ad libitum* ever since the lesion was made, with AcbC-lesioned rats weighing ~88% as much as sham-operated controls in this situation (Cardinal, 2001), much as in the present study (85%). This suggests that the food deprivation regimen maintained the proportional relationship between actual and free-feeding mass similarly in sham-operated and AcbC-lesioned rats.
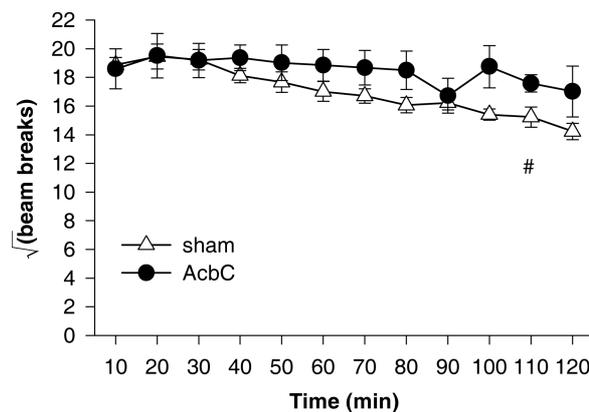
Figure 51: Locomotor activity in a novel environment in AcbC-lesioned and sham-operated rats

AcbC-lesioned rats were hyperactive compared to sham-operated controls, being slower to habituate to a novel environment. Analysis using the model lesion$_2$ × (bin$_{12}$ × S) revealed a lesion × bin interaction ($F_{8.3,133.4} = 2.20$, $\tilde{\varepsilon} = .758$, # $p = .029$), reflecting a difference in habituation between the groups, and a main effect of bin ($F_{8.34,133.4} = 9.02$, $p < .001$), reflecting habituation, though there was no main effect of lesion ($F_{1,16} = 2.24$, $p = .154$).

4.5 DISCUSSION

These results suggest that the AcbC contributes to the selection of uncertain rewards. AcbC-lesioned rats exhibited risk-averse choice: they chose large, uncertain rewards less than sham-operated controls when

offered a smaller, certain alternative, even though they showed a strong and unaltered preference for large rewards over small rewards, and for certain rewards over uncertain rewards. By the end of testing, the control group behaved as if indifferent between a single certain food pellet and four pellets delivered with $p = 0.32$ (close to the probability of 0.25 that would represent rational indifference), while the AcbC-lesioned group behaved as if indifferent between a single certain pellet and four pellets delivered with $p = 0.70$.

Though these results establish that the lesions used in this study caused this pattern of behaviour, the precise mechanism by which this occurs is unknown: for example, it is possible that the damage caused to structures adjacent to the AcbC, though limited, played a role in this pattern of choice, or that adaptations in other structures consequent upon the lesion were important in the behavioural effects (particularly given that risk aversion was not apparent immediately but emerged with further time and postoperative experience with the task).

4.5.1 Choice in normal subjects

The dominant model of uncertainty or probability discounting (Rachlin *et al.*, 1986; Rachlin *et al.*, 1991; Ho *et al.*, 1999; Green & Myerson, 2004) suggests that subjects calculate a value for each reinforcer, according to its size and other parameters, and discount this by multiplying it by $1/(1+H\theta)$, where θ represents the odds against obtaining the reinforcer, $\theta = (1 - p)/p$, and H represents an odds discounting parameter that is specific to the individual subject but stable over time for that subject. In this model, value is a hyperbolic function of the odds θ; such a hyperbolic function is supported by empirical research, at least in humans (Rachlin *et al.*, 1986; Rachlin *et al.*, 1991; Rachlin & Siegel, 1994; Kacelnik, 1997b; Richards *et al.*, 1999b; Rachlin *et al.*, 2000). The present task is not well suited to evaluating such a quantitative model, since in discrete-trial schedules it is often the case that animals maximize, or allocate most of their choices to whichever option is the more favourable (Mackintosh, 1974). However, the behaviour of normal subjects here can be evaluated as to its optimality. In the present task, neither risk aversion nor risk taking is optimal if carried to extremes. Optimal behaviour, to maximize the expected amount of food, is to choose the small/certain lever when the large (four-pellet) reinforcer probability is less than 0.25, to choose the large/uncertain lever when the probability exceeds 0.25, and to be indifferent at $p = 0.25$ (i.e. to exhibit a step function in choice). Shams' choice of the large reinforcer behaviour was better described by a linear function of the large-reinforcer probability than by such a step function. Nevertheless, shams' behaviour was reasonably close to the optimal in the most obvious way to measure optimality, namely the amount of food obtained (Figure 50b, p. 121).

4.5.2 Effects of AcbC lesions in terms of conditioning processes

AcbC-lesioned rats chose the large, uncertain reinforcer less often than shams did, but only when a smaller certain reinforcer was available as an alternative; that is, they exhibited risk-averse choice. A number of simple explanations of the present results may be ruled out. For example, it is unlikely that the pattern of choice exhibited by AcbC-lesioned rats can be explained in terms of perseveration, within a session, on the initially optimal lever. It might be that animals that perseverated on the lever delivering the small, certain reinforcer, because that lever was initially optimal, would appear to exhibit risk-averse choice in sessions in which the large-reinforcer probability increased across the session (Figure 47e, p. 115), but this could not explain the same pattern of choice in sessions in which the same lever was initially suboptimal, i.e. when the large-reinforcer probability decreased across the session (Figure 47c, p. 115). Furthermore, although AcbC lesions are known to affect processes through which Pavlovian CSs

affect behaviour, including PIT, autoshaping, and conditioned reinforcement (Parkinson *et al.*, 1999a; Parkinson *et al.*, 1999b; Parkinson *et al.*, 2000c; Hall *et al.*, 2001; Cardinal *et al.*, 2002b; de Borchgrave *et al.*, 2002; Parkinson *et al.*, 2002), there was no Pavlovian CS that was differentially associated with uncertain as opposed to certain reinforcement in this task, so these effects cannot explain the present re-sults. It might be that the AcbC lesion impaired subjects' knowledge of the instrumental action–outcome contingency specifically for the uncertain outcome. There is some debate about the role of the AcbC in instrumental conditioning (see Cardinal *et al.*, 2002a; Cardinal & Everitt, 2004; Kelley, 2004) and goal-directed action, a subset of instrumental conditioning (Dickinson, 1994; Dickinson & Balleine, 1994; Cardinal *et al.*, 2002a). Manipulation of the AcbC can certainly affect instrumental learning (Kelley *et al.*, 1997; Smith-Roe & Kelley, 2000; Baldwin *et al.*, 2002a; Hernandez *et al.*, 2002). However, the AcbC is not required for simple instrumental conditioning: rats with AcbC lesions acquire lever-press responses on FR-1 schedules at supernormal levels (Chapter 4; Cardinal & Cheung, 2005), and rats with Acb or AcbC lesions are fully sensitive to changes in the action–outcome contingency (Balleine & Killcross, 1994; Corbit *et al.*, 2001; de Borchgrave *et al.*, 2002). However, when acquiring a sequence of RR sched-ules, AcbC-lesioned rats respond somewhat less than sham-operated controls (Corbit *et al.*, 2001), while lesions of the whole Acb made rats respond slightly, though not significantly, less on a similar sequence of RR schedules (de Borchgrave *et al.*, 2002). RR schedules clearly involve probabilistic reinforcement, so these results are consistent with the possibility that the present impairment shown by AcbC-lesioned rats in choosing large, unlikely rewards is due to impaired instrumental conditioning when the outcome is uncertain—and, conversely, that the impairment in simple instrumental learning seen previously (Corbit *et al.*, 2001) was specifically a result of the reward uncertainty inherent in a RR schedule, given that AcbC-lesioned rats learn instrumental responses normally or supernormally with certain immediate rein-forcement (Chapter 4; Cardinal & Cheung, 2005). It is also possible that AcbC-lesioned rats represent the instrumental contingency normally with uncertain reward, but simply value the uncertain outcome less and respond less for it accordingly, as discussed next.

4.5.3 Effects of AcbC lesions in terms of probability discounting and reinforcer magnitude sensitivity

Since the present study required rats to choose between small, certain and large, uncertain rewards, an effect of the lesion to alter the perception of relative reward magnitude might affect choice, just as an al-teration in the perception of reward probability might. For example, altering the absolute magnitudes of the reinforcers can affect choice involving probabilistic reinforcement (Mazur, 1988; Kirby & Marakovic, 1996), as would be predicted if reinforcer "value" is not simply a linear function of physical magnitude (Ho *et al.*, 1999). Specifically, the present results (a tendency for AcbC-lesioned rats to choose the small, certain reinforcer more than shams) could be explained by "risk aversion" (increased or steeper uncer-tainty/odds/probability discounting), or if the difference between 1 and 4 pellets was perceived to be smaller by AcbC-lesioned subjects than by shams (due to reduced discrimination between the two rein-forcer magnitudes, or perhaps with a normal ability to tell the two apart but with an altered perception of relative value). For example, if a normal subject assigned values of 1 and 4 to the reinforcers, and a le-sioned subject assigned values of 1 and 3 to the same reinforcers, then the lesioned subject would be less likely than the sham to choose the large reinforcer when it was made uncertain, even without any primary abnormality in the processing of probability. At first glance, this interpretation would appear to be sup-ported by the observation that AcbC-lesioned rats chose the large reinforcer somewhat less often than shams when it was certain, as well as when it was uncertain. However, several lines of evidence suggest

this explanation is not the correct one. When the large and the small reinforcers were both made consistently certain, there were no differences between AcbC-lesioned rats and controls (Figure 47d, p. 115). Furthermore, other evidence indicates that AcbC lesions do not impair reinforcer magnitude discrimination or the perception of relative reinforcer value. Excitotoxic lesions of the whole Acb do not prevent rats from detecting changes in reward value, induced either by altering the concentration of a sucrose reward or by changing the deprivational state of the subject (Balleine & Killcross, 1994). Such lesions also do not impair rats' ability to respond faster when environmental cues predict the availability of larger rewards (Brown & Bowman, 1995), and nor does inactivation of the Acb with local anaesthetic or blockade of AMPA glutamate receptors in the Acb (Giertler *et al.*, 2004; 2005); the effects of intra-Acb NMDA receptor antagonists have varied (Hauber *et al.*, 2000; Giertler *et al.*, 2003; 2005). AcbC-lesioned rats can still discriminate large from small rewards (Cardinal *et al.*, 2003b; 2004). Similarly, DA depletion of the Acb does not affect the ability to discriminate large from small reinforcers (Salamone *et al.*, 1994; Cousins *et al.*, 1996; Salamone *et al.*, 2001), and systemic DA antagonists do not affect the perceived quantity of food as assessed in a psychophysical procedure (Martin-Iverson *et al.*, 1987). Furthermore, a recent study found evidence that AcbC-lesioned rats may even show somewhat enhanced reinforcer magnitude discrimination, or an exaggerated perception of relative value (Chapter 4; Cardinal & Cheung, 2005). Given that reinforcer magnitude discrimination appears to be unimpaired, at worst, by AcbC lesions, the observation in the present study that AcbC-lesioned rats chose the large reinforcer somewhat less often than controls in the task in which large-reinforcer probabilities changed throughout the session is more likely to be explained by within-session generalization (Evenden & Ryan, 1996; Cardinal *et al.*, 2000; 2003b)—i.e. that avoidance of the large reinforcer during trial blocks when it was uncertain generalized to trial blocks when it was certain. Together, these findings suggest that the present results are best explained as an effect of AcbC lesions to increase the rate of uncertainty/odds/probability discounting—effectively, a tendency to behave as if an uncertain outcome were less likely than it really is.

4.5.4 Probability versus delay discounting

It is known that AcbC lesions affect choice and learning involving delayed reinforcement (Chapters 1 and 4; Cardinal *et al.*, 2001; 2003b; Cardinal & Cheung, 2005). As discussed earlier, it has been suggested that delay or temporal discounting, the process by which delayed reinforcers lose value, and probability or odds discounting, the process by which uncertain reinforcers lose value, reflect the same underlying process (Rachlin *et al.*, 1986; Stevenson, 1986; Rachlin *et al.*, 1987; Mazur, 1989; Rachlin *et al.*, 1991; Mazur, 1995; Green & Myerson, 1996; Mazur, 1997; Sozou, 1998). For example, in the present task, choosing the uncertain large reinforcer five times but only obtaining it on the fifth response might be seen as equivalent to a very long delay, on average, between choice of the large reinforcer and its eventual delivery. Alternatively, delays may be seen as entailing the ecological risk of losing the reward during the delay. The failure of AcbC-lesioned rats to choose an uncertain reinforcer (risk aversion, as seen in the stable phase of the present results) and their failure to choose a delayed reinforcer may therefore be explained in the same way. However, there is evidence that time and probability discounting are different and dissociable processes (Ho *et al.*, 1999; Mitchell, 2003; Green & Myerson, 2004). Most simply, it is not surprising that currency inflation affects human decisions involving delayed but not probabilistic financial reward (Ostaszewski *et al.*, 1998). Moreover, the absolute magnitude of rewards can have different effects on delayed and probabilistic discounting (Green *et al.*, 1999; Myerson *et al.*, 2003; Green & Myerson, 2004). A study looking at human choices in a gambling task found that individuals' propensity to choose rapidly (one, perhaps motoric, measure of delay aversion) and their propensity to bet large

amounts of money on uncertain outcomes (a measure of risk taking) represented independent factors (Deakin *et al.*, 2004). Some studies have found abnormal delay discounting, but not uncertainty discounting, in drug addicts (Vuchinich & Calamas, 1997; Mitchell, 1999; Mitchell, 2003; Reynolds *et al.*, 2004b), while gamblers have been observed to discount probabilistic rewards less steeply than controls (i.e. to take risks) without showing differences in delay discounting (Holt *et al.*, 2003).

4.5.5 Implications for AcbC function and impulsivity

Impulsivity is multifaceted, reflecting—at the least—individual differences in distinct and dissociable processes involving information gathering, the selection of outcomes, and the inhibition of motor actions (Evenden, 1999b). Furthermore, as discussed above, delay discounting and probability discounting may also reflect separate processes. Damage to the AcbC can produce impulsive choice in the sense of an impaired ability to choose delayed rewards (Cardinal *et al.*, 2001), in addition to hyperactivity (Chapter 4; Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001; Cardinal & Cheung, 2005), though without impairments in attentional function (Christakou *et al.*, 2004) and without motoric impulsivity as assessed by the stop-signal task (Eagle & Robbins, 2003). In the context of choice involving uncertain appetitive reinforcement, "impulsivity" would equate to risk taking—less steep uncertainty discounting, or greater willingness to choose unlikely rewards. AcbC lesions, however, produced a risk-averse or conservative pattern of choice in the present study. Clearly, then, AcbC-lesioned rats cannot be characterized as impulsive in all senses. A more appropriate unifying concept would seem to be that the AcbC promotes the selection, and perhaps the salience, of uncertain and delayed rewards—perhaps, in general, of rewards that are not certain, imminent, or present (Cardinal *et al.*, 2002a). The AcbC promotes choice of (Cardinal *et al.*, 2001) and learning with (Chapter 4; Cardinal & Cheung, 2005) delayed rewards. It appears to promote the selection of uncertain reinforcers (present results), and this is compatible with human imaging studies showing increased Acb activation during the selection of high-risk options (Ernst *et al.*, 2004; Matthews *et al.*, 2004). The Acb is required for PIT, the process by which Pavlovian CSs signalling reward enhance instrumental responding for those rewards (Hall *et al.*, 2001; de Borchgrave *et al.*, 2002). It is also required for autoshaping, or locomotor approach to appetitive Pavlovian CSs (Parkinson *et al.*, 1999a; Parkinson *et al.*, 1999b; Parkinson *et al.*, 2000c; Cardinal *et al.*, 2002b; Parkinson *et al.*, 2002), and it influences conditioned reinforcement, the process of working for CSs previously paired with reinforcement (Taylor & Robbins, 1984; 1986; Cador *et al.*, 1991; Parkinson *et al.*, 1999a). Acb DA also contributes to subjects' motivation to work hard (Ikemoto & Panksepp, 1999; Salamone & Correa, 2002; Salamone *et al.*, 2003; Mingote *et al.*, 2005; Salamone *et al.*, 2005).

It is not known whether AcbC lesions would produce similar effects on choice involving uncertain aversive events. It would be expected that increased odds/uncertainty/probability discounting—effectively, a tendency to behave as if an uncertain outcome were *less likely* than it really is—would produce risk aversion for appetitive outcomes (reduced willingness to choose large, unlikely rewards) but risk proneness for aversive outcomes (increased willingness to choose large, uncertain punishments over small, certain punishments) (Ho *et al.*, 1999). In humans, at least, the delay and probability discounting processes appear similar for rewards and losses (Ostaszewski & Karzel, 2002; Green & Myerson, 2004).

4.5.6 Relationship to structures and neuromodulator systems innervating the AcbC

The PFC, which projects heavily to the AcbC (Brog *et al.*, 1993), is also involved in decision making under conditions of uncertainty. Humans with OFC or ventromedial PFC damage are impaired on the Iowa gambling task (Bechara *et al.*, 1994; 1996; 1997), in which subjects must learn to differentiate between

low-reward, low-risk card decks that yield a net positive outcome and high-reward, high-risk decks that yield a net negative outcome, though the precise locus and nature of the deficit seen on this task is debated (Manes *et al.*, 2002; Clark *et al.*, 2003; Fellows & Farah, 2005). Choice between small, likely rewards and large, unlikely rewards increases cerebral blood flow and BOLD signal in orbital and inferior PFC (Rogers *et al.*, 1999b; 2004b), and OFC damage also impairs performance of a task requiring human subjects to choose between two possible outcomes and to bet on their choice, with lesioned subjects deciding slowly and failing to choose the optimal, most likely outcome (Rogers *et al.*, 1999a). Excitotoxic lesions of the OFC make rats less likely than sham-operated controls to choose a large, uncertain reward over a small, certain reward (Mobini *et al.*, 2002); OFC-lesioned rats had lower indifference odds (higher indifference probabilities; steeper uncertainty discounting) and exhibited risk-averse choice, just like the AcbC-lesioned subjects in the present study. There is direct evidence that OFC lesions do alter sensitivity to the relative magnitudes of the two rewards (Kheramin *et al.*, 2005), as does OFC DA depletion (Kheramin *et al.*, 2004), but the effects on uncertainty discounting are present in addition to those on reinforcer magnitude sensitivity (Kheramin *et al.*, 2003).

The Acb is also innervated by the a number of neuromodulator systems, including the 5-HT system (Halliday *et al.*, 1995). Although manipulations of 5-HT influence choice involving delayed reinforcement, there is less evidence that they influence choice involving uncertainty and risk. Correlational studies have indicated that low CSF levels of the 5-HT metabolite 5-HIAA are associated with risk taking in monkeys (Mehlman *et al.*, 1994) and impulsive aggression, violence, and suicide in humans (Åsberg *et al.*, 1976; Linnoila *et al.*, 1983; Brown & Linnoila, 1990; Linnoila *et al.*, 1993; Mann, 2003). Forebrain 5-HT depletion tends to steepen temporal (delay) discounting (reviewed briefly by Cardinal *et al.*, 2004); however, it does not appear to influence choice involving probabilistic reinforcement. Dietary tryptophan depletion (Biggio *et al.*, 1974; Clemens *et al.*, 1980; Delgado *et al.*, 1989) decreases levels of 5-HT metabolites in CSF, an indirect indicator of brain 5-HT levels, but has not been shown to affect probability discounting in humans (Anderson *et al.*, 2003; Rogers *et al.*, 2003); similarly, forebrain 5-HT depletion in rats does not affect choice between small, certain rewards and large, uncertain rewards (Mobini *et al.*, 2000b). The AcbC also receives a substantial DA innervation, and DA neurons respond to reward prediction errors (Schultz *et al.*, 1997; Schultz, 1998; Schultz *et al.*, 1998; Schultz & Dickinson, 2000; Schultz, 2006). Although systemic D_2-type DA receptor antagonists can induce impulsive choice involving delayed reinforcement (Wade *et al.*, 2000), this effect may not occur in the Acb (Winstanley *et al.*, 2005b), the response of DA neurons specifically to reward uncertainty is debated (Fiorillo *et al.*, 2003; Niv *et al.*, 2005; Tobler *et al.*, 2005), and little is known of the role of DA in choice involving uncertain rewards. Systemic noradrenergic (NA) blockade has also been shown to affect decision making under uncertainty in humans, by reducing the discrimination between magnitudes of different losses when the probability of losing was high (Rogers *et al.*, 2004a), though NA reuptake inhibition has not been shown to affect the Iowa gambling task (O'Carroll & Papps, 2003). However, the Acb does not receive a substantial NA innervation (Aston-Jones *et al.*, 1995).

4.6 CONCLUSIONS

These experiments have shown that excitotoxic lesions of the AcbC induce risk-averse choice in rats. AcbC lesions did not prevent rats from discriminating a large reward from a small reward, or a certain reward from an uncertain reward. However, when offered the choice between a small/certain reward and

a large/uncertain reward, AcbC-lesioned rats showed a reduced preference for the large/uncertain reward (compared to sham-operated controls) in their final pattern of postoperative choice. AcbC-lesioned rats exhibited a tendency to behave as if an uncertain outcome were less likely than was really the case. Together with previous studies, these results suggest that the AcbC contributes to reinforcement and choice particularly when the reinforcer is temporally distant or uncertain.

Chapter 5: General discussion

5.1 OVERVIEW

The experiments described in this thesis addressed the role played by the AcbC in rats' ability to learn from delayed rewards, to perform previously learned actions for delayed rewards, to assess reward magnitudes, and to choose uncertain rewards, together with the role of the hippocampus in the ability to learn from delayed rewards and to choose delayed rewards. In this concluding chapter, the findings from these experiments will first be summarized briefly. The results have already been discussed in Chapters 2–4; in this chapter, their implications will be considered in the wider context of impulse control disorders and the neural mechanisms that underlie reinforcement.

5.2 SUMMARY OF RESULTS

5.2.1 Role of the AcbC in learning with delayed reward

In Chapter 2 (Cardinal & Cheung, 2005), it was shown that excitotoxic lesions of the AcbC did not prevent rats from learning a simple instrumental response when the reinforcing outcome followed their action immediately. However, AcbC lesions impaired rats' ability to learn the same instrumental response when the outcome was delayed by 10 or 20 s. Increasing delays impaired learning in normal rats to some degree, which is a well-known finding (Grice, 1948; Lattal & Gleeson, 1990; Dickinson *et al.*, 1992). Rats with AcbC lesions were unimpaired (compared to sham-operated controls) when there was no delay, but were profoundly impaired when there was a delay between action and outcome, compared to shams learning with the same delay. AcbC lesions also impaired performance of an instrumental response that was learned preoperatively, but again only when response–reinforcer delays were present.

The fact that pre-exposure to the context improves instrumental learning in normal rats (Dickinson *et al.*, 1992) suggests one possible mechanism by which AcbC lesions might retard learning when delays are present. When a reinforcer arrives, it may be associated either with a preceding response, or with the context. Therefore, in normal animals, pre-exposure to the context may retard the formation of context–reinforcer associations by latent inhibition, or it might serve to retard the formation of associations between irrelevant behaviours and reinforcement. Non-reinforced exposure to the context forces the subjects to experience a zero-response, zero-reinforcer situation, i.e. $P(outcome \mid no\ action) = 0$. When they are then exposed to the instrumental contingency, such that $P(outcome \mid action) > 0$, this prior experience may enhance their ability to detect the instrumental contingency $\Delta P = P(outcome \mid action) - P(outcome \mid no\ action)$. In one aversive Pavlovian conditioning procedure in which a CS was paired with electric shock, AcbC lesions have been shown to impair conditioning to discrete CSs, but simultaneously to enhance conditioning to contextual or background CSs (Parkinson *et al.*, 1999b), though not all behavioural paradigms show this effect (Levita *et al.*, 2002; Jongen-Relo *et al.*, 2003). It is therefore possible that enhanced formation of context–reinforcer associations may explain the retardation of response–reinforcer learning in AcbC-lesioned rats in the presence of delays.

Acb lesions have also produced delay-dependent impairments in a delayed-matching-to-position task (Dunnett, 1990; Reading & Dunnett, 1991). Their effects on the delayed-matching-to-sample paradigm

have also been studied, but a more profound and delay-independent deficit was observed, likely due to differences in the specific task used (Burk & Mair, 2001).

5.2.2 Role of the AcbC in assessing reward magnitude

Previous studies have found that excitotoxic lesions of the whole Acb do not prevent rats from detecting changes in reward value (Balleine & Killcross, 1994). Such lesions also do not impair rats' ability to respond faster when environmental cues predict the availability of larger rewards (Brown & Bowman, 1995), and nor does inactivation of the Acb with local anaesthetic or blockade of AMPA glutamate receptors in the Acb (Giertler *et al.*, 2004). The effects of intra-Acb NMDA receptor antagonists have varied (Hauber *et al.*, 2000; Giertler *et al.*, 2003). AcbC-lesioned rats can still discriminate large from small rewards (Cardinal *et al.*, 2003b; 2004). Similarly, DA depletion of the Acb does not affect the ability to discriminate large from small reinforcers (Salamone *et al.*, 1994; Cousins *et al.*, 1996; Salamone *et al.*, 2001), and systemic DA antagonists do not affect the perceived quantity of food as assessed in a psychophysical procedure (Martin-Iverson *et al.*, 1987). These studies suggest that AcbC lesions do not prevent rats from discriminating *qualitatively* between large and small rewards, and that DA antagonism does not alter quantitative reward magnitude discrimination. For the purposes of analyses involving reward delay, it is important to know whether AcbC lesions alter the *quantitative* perception of reward magnitude—e.g. whether such lesions alter the magnitude sensitivity parameter Q in the model of Ho *et al.* (1999). In Chapter 2 (Cardinal & Cheung, 2005), it was observed that excitotoxic AcbC lesions did not impair, but rather improved, rats' ability to allocate their responses across two schedules in proportion to the experienced reinforcement rate, even when the two schedules were identical except in the magnitude of the reinforcements they provide, suggesting their sensitivity to reinforcer magnitude is quantitatively no worse than shams'.

5.2.3 Role of the hippocampus in learning with and choosing delayed reward

As discussed in Chapter 1, a role of the hippocampus in learning with delayed reinforcement might be suspected, because there is good evidence that the hippocampus contributes to the representation of context (Hirsh, 1974; Good & Honey, 1991; Selden *et al.*, 1991; Kim & Fanselow, 1992; Phillips & LeDoux, 1992; Honey & Good, 1993; Jarrard, 1993; Kim *et al.*, 1993; Phillips & LeDoux, 1994; Phillips & LeDoux, 1995; Chen *et al.*, 1996; Maren & Fanselow, 1997; Anagnostaras *et al.*, 1999; Holland & Bouton, 1999; Good, 2002; Rudy *et al.*, 2002; Ito *et al.*, 2005) and, as discussed earlier, contextual conditioning is important in learning with delays. Since context–outcome associations are thought to hinder instrumental learning with delayed reinforcement through contextual competition (Dickinson *et al.*, 1992; Dickinson & Balleine, 1994), it follows that if H lesions impair the formation of associations involving the context, such lesions might reduce contextual competition and hence facilitate instrumental conditioning when there is an action–outcome delay.

Indeed, excitotoxic lesions of the H ameliorated the deleterious effects of response–reinforcer delays on instrumental learning (Chapter 3; Cheung & Cardinal, 2005). H-lesioned rats responded slightly less than controls in the absence of delays, but they became better at learning (relative to shams) as the delays increased, in a delay-dependent fashion. This may have been because the lesion hindered the formation of context–outcome associations, promoting response–outcome association instead.

Unexpectedly, in separate experiments H-lesioned rats exhibited impulsive choice, preferring an immediate, small reward to a delayed, larger reward (in a task based on that of Evenden & Ryan, 1996), even though they preferred the large reward when it was not delayed (Chapter 3; Cheung & Cardinal,

2005). Though a quantitative difference in sensitivity to reinforcer magnitude might explain these results, as discussed above (Ho *et al.*, 1999) (see Chapter 1, p. 40), H-lesioned rats were able to discriminate the large from the small reinforcer, and such evidence as exists suggests that H-lesioned rats perceive reward magnitude normally (Kesner & Williams, 1995; Gilbert & Kesner, 2002). These results may also be explained in terms of altered temporal perception (as discussed on p. 106), affecting choice prospectively or retrospectively. For example, a lesion that increased the speed of an "internal clock" (Gibbon *et al.*, 1997) might affect choice prospectively in this task (i.e. the lesioned subject perceives itself to be at a later time point in the session than it actually is; since the task used a delay for the LL reward that increased across the session, such an effect would hasten the within-session shift towards the SS alternative), or might affect retrospective choice (i.e. the lesioned subject experiences the delay to the large reinforcer as longer than it actually is, causing it to value the reinforcer less than shams). The evidence for the role of the hippocampus in temporal perception is inconclusive: some studies have found that aspirative hippocampal lesions did not affect timing behaviour (Rawlins *et al.*, 1983; Port *et al.*, 1986; Dietrich *et al.*, 1997; Dietrich & Allen, 1998), whereas others have suggested that lesions of the hippocampus or fimbria/fornix speed up an internal clock, or reduce the estimation of time periods when a stimulus being timed is interrupted (Meck *et al.*, 1984; Olton *et al.*, 1987; Meck, 1988; Hata & Okaichi, 1998; Wallenstein *et al.*, 1998). In any case, H-lesioned rats were better at learning with delayed reinforcement but worse at choosing it, suggesting that self-controlled choice and learning with delayed reinforcement tax different psychological processes.

5.2.4 Role of the AcbC in choosing uncertain reward

As discussed in Chapter 1, correlational studies have suggested that the Acb may also be involved in the processing of uncertain or probabilistic reinforcement (Fiorillo *et al.*, 2003; Aron *et al.*, 2004; Ernst *et al.*, 2004; Matthews *et al.*, 2004; Fiorillo *et al.*, 2005; Niv *et al.*, 2005; Tobler *et al.*, 2005), yet this issue had not previously been addressed in a controlled interventional study. In Chapter 4 (Cardinal & Howes, 2005), excitotoxic lesions of the AcbC were found to induce what might be characterized as risk-averse choice in rats. AcbC lesions did not prevent rats from discriminating a large reward from a small reward, or a certain reward from an uncertain reward. However, when offered the choice between a small/certain reward and a large/uncertain reward, AcbC-lesioned rats showed a reduced preference for the large/uncertain reward (compared to sham-operated controls) in their final pattern of postoperative choice. AcbC-lesioned rats exhibited a tendency to behave as if an uncertain outcome were less likely than was really the case. Together with studies examining the effects of AcbC lesions on delayed reinforcement, these results suggest that the AcbC contributes to reinforcement and choice particularly when the reinforcer is temporally distant or uncertain.

5.3 WIDER IMPLICATIONS

There is evidence that the AcbC is involved in the pathogenesis of impulsive choice: the integrity of the AcbC is critical for animals to tolerate delays to appetitive reinforcement (Cardinal *et al.*, 2001), to learn normally from delayed appetitive reinforcement (Cardinal & Cheung, 2005), and to choose uncertain appetitive reinforcement normally (Cardinal & Howes, 2005). Likewise, normal hippocampal function appears necessary for rats to choose delayed appetitive reinforcement normally (Cheung & Cardinal, 2005), although the hippocampus appears to make a different contribution to learning with delayed reinforcement. In addition to providing neuroanatomical insight into the normal process through which delayed and/or uncertain reinforcement affects behaviour, this finding suggests a mechanism by which dysfunc-

tion of these structures may contribute to addiction, ADHD, and other impulse control disorders. In this section, I will set the contribution of the hippocampus to delayed reinforcement in the wider context of theories of time-limited hippocampal memory storage, and discuss broader implications of the present findings regarding hippocampal and AcbC function for disorders of impulse control.

5.3.1 The hippocampus and time-limited memory storage

The involvement of the hippocampus in learning with delayed reinforcement was hypothesized (Chapter 1, p. 49; Chapter 3, p. 81) to be a consequence of its role in representing contexts. In turn, this may be due to the ability of the hippocampus rapidly to associate arbitrary stimuli (see Chapter 1, p. 34).

It is not known whether the delay-related contribution of the hippocampus to instrumental learning is permanent or transient. However, many types of memory that are initially dependent upon the hippocampus do not remain so. This concept originally stemmed from the observation of a temporally graded retrograde amnesia in humans following medial temporal lobe resection or more restricted hippocampal damage, with good memory for events long past but poor memory for relatively recent events preceding the insult, in addition to the more obvious profound anterograde amnesia (Scoville & Milner, 1957; Zola-Morgan *et al.*, 1986; Corkin *et al.*, 1997; Corkin, 2002). Similar effects were observed in many animal models involving lesions restricted to the hippocampal formation (see Squire *et al.*, 2001). These observations led to the hypothesis that the hippocampus is involved in consolidating memories held elsewhere (Scoville & Milner, 1957; Squire *et al.*, 1975; 1980; Squire, 1986; 1992; Squire *et al.*, 2001): recent memories are vulnerable to hippocampal damage, but with time they become independent of the hippocampus, perhaps depending instead on cortical sites.

The major competing view is the "multiple memory trace" hypothesis of Nadel & Moscovitch (1997). They argue that the duration of retrograde amnesia for human autobiographical episodes following medial temporal lobe damage is extremely long (25–40 years), implying that most humans throughout history would never have "fully" consolidated a memory, and that the retrograde amnesia may not even be temporally graded at all (i.e. that the hippocampus causes a "flat" retrograde amnesia, with loss of all memories that ever depended upon it). Nadel & Moscovitch (1997) consider the hippocampus to be permanently involved in the storage of autobiographical memories, taking the viewpoint that autobiographical memory, personal semantic memory, and "general" semantic memory (vocabulary, grammar, object recognition) are progressively less sensitive, in that order, to retrograde amnesia following medial temporal lobe lesions in humans. In their view, the hippocampus provides a permanent spatial or contextual "index" that helps to retrieve a given memory. One-off (e.g. recent) autobiographical memories are dependent upon their index for retrieval, so are vulnerable to hippocampal damage. Semantic information is extracted from repeated episodic experiences; therefore, semantic information—and well-rehearsed, i.e. old, autobiographical memory—is supported by multiple memory traces, and is less dependent upon the hippocampal "contextual index" for retrieval. Recent statements of this hypothesis have been provided by Nadel & Bohbot (2001) and Rosenbaum *et al.* (2001).

However, retrograde amnesia is difficult to study in humans, because it is necessarily done retrospectively—the experimenter must assess the subject's memory for recent and ancient experience after the onset of amnesia, but it is difficult to sample memory equivalently from different past time periods, and to know that these memories were of comparable "strength" before the event that caused amnesia. The ideal test to compare these two hypotheses therefore involves prospective studies in animals (see Murray & Bussey, 2001, for these and other important methodological issues). The majority of such studies have shown temporally graded retrograde amnesia following a variety of hippocampus, fornix, and entorhinal

cortex lesions (see Squire *et al.*, 2001), supporting the view that the hippocampus does play a transient role in the storage of at least some types of memory. Amongst these studies, electrolytic or excitotoxic lesions of the hippocampus produce a time-limited retrograde amnesia for contextually conditioned fear (see Anagnostaras *et al.*, 2001).

Thus, memories of certain kinds initially depend upon the hippocampus but with time they become independent of the hippocampus. Transient hippocampal involvement does not require memories to "move" in a physically arbitrary way; there are perfectly plausible ways in which a memory might depend on a structure only temporarily (e.g. McClelland *et al.*, 1995). Figure 52 illustrates one possible simple mechanism. Recent studies have provided direct support for the view that hippocampal–cortical interactions are involved in the consolidation of some types of memory (Maviel *et al.*, 2004).
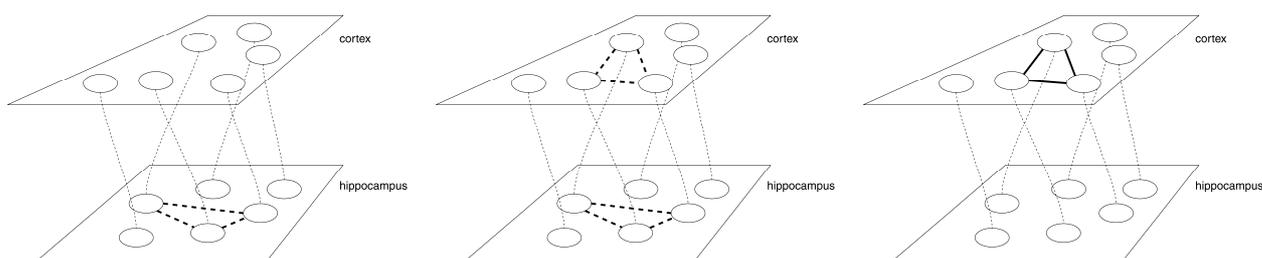
Figure 52: A simple mechanism for transient involvement of the hippocampus in memory storage
Left to right: schematics of how the hippocampus might interact with the cerebral cortex to consolidate memories "held" elsewhere. If the hippocampus exhibits rapid synaptic plasticity (but this is transient or easily disrupted) and the cortex exhibits slower but more stable plasticity, a plausible mechanism might proceed as follows. **Left:** hippocampal neurons have permanent connections to regions of neocortex (vertical dotted lines). A memory is formed by the hippocampus rapidly associating a number of active neurons, via synaptic plasticity (horizontal dashed lines). The memory is dependent upon the hippocampus. **Centre:** subsequent hippocampal activity promotes the firing of a cortical network that corresponds to the group of associated hippocampal neurons. As a direct result, this promotes an increase in the connectivity between the cortical neurons. **Right:** with time, the cortical links become strong enough not to require further hippocampus-driven consolidation. The memory is now independent of the hippocampus.

The impermanence of hippocampal memories has been demonstrated both at the behavioural and the synaptic level. Active processes appear to be involved in the decay of hippocampal memories. For example, Villarreal *et al.* (2002) have shown that systemic administration of the NMDA antagonist 3-(2-carboxypiperazin-4-yl)-propyl-1-phosphonic acid (CPP) blocks decay of hippocampal LTP; when given systemically between training and testing of performance in a radial 8-arm maze task known to depend on the hippocampus, CPP improved the retention of the memory, though it was not shown that this was due specifically to the drug's effect on the hippocampus. Perhaps decay of LTP, or long-term depression (LTD), which also depends on NMDA receptors (Dudek & Bear, 1992; Morris, 1994), is required to allow the hippocampus to acquire new memories, at the expense of old ones. If a rapidly associating network does not have the ability to lose old memories, there is catastrophic interference when new memories are laid down; this is the stability–plasticity dilemma familiar to connectionist modellers (Grossberg, 1982; McCloskey & Cohen, 1989). Rosenzweig *et al.* (2002) suggest that Villarreal *et al.* (2002) blocked exactly this loss of old memories with CPP.

Moreover, old memories that were once dependent upon the hippocampus can become so again. A "standard" view of consolidation would be that memories are created in a labile state (sometimes thought of as short-term memory, although this term has other uses), and with time, they are consolidated into a stable state (sometimes termed long-term memory). For example, electroconvulsive shock (ECS) or electroconvulsive therapy (ECT), which disrupts all ongoing electrical activity in the regions of the brain to

which current is applied, induces amnesia if given shortly after training, but not if given a long time after training (Duncan, 1949; Squire *et al.*, 1975). While the formation of new memories does not require protein synthesis, the consolidation of memories does; thus, administering the protein synthesis inhibitor anisomycin during contextual fear conditioning does not impair the memory of mice if they are tested one hour later, but that memory fades by 24 h as compared to a control group (see e.g. Abel *et al.*, 1997; Kandel, 2001). The same is true of hippocampal LTP: "early" LTP is not dependent upon protein synthesis, but it fades; normally, it is made long lasting by a second phase, "late" LTP, which requires protein synthesis (see Beggs *et al.*, 1999).

This view is extended by the concept of reconsolidation. As before, this hypothesis suggests that memories are created in a labile state and are consolidated into a stable state. However, in this theory, recalling or reactivating a memory *returns it to the labile state*. Therefore, although protein synthesis inhibitors or other amnestic treatments do not disrupt stable memories, they should be able to disrupt old memories that have been reactivated. Indeed, this has been observed (Misanin *et al.*, 1968). Recently, Nader *et al.* (2000) found that infusions of anisomycin into the BLA, a critical site of plasticity for CS–US associations involved in conditioned freezing in the rat, disrupted memory for a CS–US association that had been "retrieved" by presenting the CS. This disruption did not occur if anisomycin was given without representation of the CS. The molecular mechanisms of consolidation and reconsolidation are doubly dissociable (Lee *et al.*, 2004), so they are not exactly the same process. Reconsolidation is receiving considerable attention at the moment (Nader, 2003), partly because of the obvious clinical potential for selective memory "erasure"; if this were achieved safely it would have enormous implications for disorders in which aberrant memories play a prominent role, including obsessive–compulsive disorder (OCD), drug addiction, post-traumatic stress disorder, and so on. To date, few clinical studies have been based on the principle of reconsolidation. One notable exception is a series by Rubin *et al.* (1969; Rubin, 1976), who gave ECT to patients with OCD after reactivating their problematic compulsion, with apparently considerable success relative to conventional, non-reactivation ECT under anaesthetic. Some cautions have been raised, not all of them critical for the clinical implications; for example, some of the effects attributed to inhibition of protein synthesis have on occasion turned out to be due to unrelated side effects of particular drugs, with these side effects affecting consolidation or retrieval (Flexner *et al.*, 1963; 1967; Davis & Squire, 1984). Likewise, it has been a matter of enduring debate whether amnesia is a result of a storage deficit or a retrieval deficit (e.g. Warrington & Weiskrantz, 1970; Squire, 1980; Squire *et al.*, 1987). Many forms of amnesia can be reversed by reminder treatments, indicating that the memories were present all along and the deficit was one of retrieval (Millin *et al.*, 2001). Typical animal studies used ECS to induce amnesia; subsequent exposure to the CS, the US, or the ECS have all been shown to reverse the amnesia (Miller & Springer, 1972; Springer & Miller, 1972; Miller *et al.*, 1974; see Millin *et al.*, 2001). This applies equally to reconsolidation studies (Millin *et al.*, 2001): again, "reminder" effects occur, implying a retrieval deficit (Judge & Quartermain, 1982; Mactutus *et al.*, 1982; Debiec *et al.*, 2002).

This reconsolidation phenomenon has been termed "cellular reconsolidation", in which reactivation of a memory returns it to a labile state at the same neural site. A further phenomenon is "systems reconsolidation" (Debiec *et al.*, 2002), in which reactivation of a memory appears to make the memory depend upon a structure that it once depended upon before. Debiec *et al.* gave rats CS(context)–US(shock) pairings. Such associations are known to depend on the hippocampus early after learning, but with consolidation they become independent of the hippocampus (see above). After 45 days, they then presented the CS on its own (or not) and lesioned the hippocampus (or not). In the absence of CS presentation, the memory did not depend on the hippocampus (no effect of the lesion); presentation of the CS caused the memory to

depend on the hippocampus again, but only for ~48 hours. Debiec *et al.* suggest, based on these and other experiments, that a memory is formed, initially depends on the hippocampus, and during this time it can undergo "cellular" reconsolidation if activated. With time, the memory is consolidated in neocortex and no longer requires the hippocampus, unless it is reactivated, in which case it depends on the hippocampus for a while (albeit for a shorter time than during initial consolidation), and so on.

These time-limited memory storage phenomena are indirectly relevant to the issues of instrumental free-operant learning with delayed reinforcement and impulsive choice, though not specifically to their relationship with contextual conditioning (discussed in Chapter 3, p. 100). It is not known how the hippocampus contributes to performance of instrumental responses learned with delayed reinforcement. Hippocampal lesions made before training delay-dependently improved free-operant instrumental learning with delayed reinforcement (Figure 40, p. 92; Figure 42, p. 95) in that delays retarded learning less in H-lesioned subjects than in shams. If this was due to a hippocampus-dependent contextual memory competing with the instrumental response for association with the reinforcer, then since one would expect long-established contextual memories to have become relatively independent of the hippocampus (Anagnostaras *et al.*, 2001), it may be that pre-exposure to the experimental context, in addition to improving learning itself (Dickinson *et al.*, 1992), would reduce the effect of hippocampal lesions made before the instrumental learning task. Potentially, by the systems reconsolidation argument, contextual retrieval might increase the effects of hippocampal lesions again. It is more difficult to predict what would happen if hippocampal lesions were made after training on this task. If a contextual representation competes during *performance*, as well as learning, of an instrumental response, and the hippocampal lesion were made whilst that contextual memory was still dependent upon the hippocampus, then one would expect hippocampal lesions to produce a delay-dependent improvement in performance of a previously learned instrumental response, in addition to any delay-independent effects. The hippocampus appears to play a time-limited role in trace eyeblink conditioning (Takehara *et al.*, 2002); however, as discussed on pp. 100 and 103, the effects of hippocampal lesions on trace conditioning and instrumental conditioning with delayed reinforcement differ even when the lesions are made before training, and the conceptual relationship between the two tasks is not perfectly clear.

In the case of the impulsive choice task (Figure 43, p. 97), hippocampal lesions were made after 19 sessions of training on the task, when performance was stable and subjects were well trained. Since H lesions impaired subjects' ability to choose the large delayed reinforcer, this suggests (but does not demonstrate conclusively) that the hippocampus makes an enduring contribution to promoting the choice of delayed reinforcers. Finally, it is not known what the effects of hippocampal lesions made prior to training on this task would be. One would expect three competing effects: a relative delay-dependent enhancement of learning the action–outcome contingency for delayed reinforcement, a retardation of learning of this contingency at zero delay, and a reduction in preference for the delayed reinforcer.

5.3.2 ADHD

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. This may make it possible to distinguish the brain regions that underlie different types of impulsivity (Evenden, 1999b), and to segregate the neural abnormalities that contribute to complex disorders such as ADHD and drug addiction, as well as to normal variation in impulsive behaviour such as during adolescence (discussed later).

The integrity of the Acb is critical for animals to tolerate delays to appetitive reinforcement (Cardinal

et al., 2001; Cardinal & Cheung, 2005). In addition to being impulsive, AcbC-lesioned rats are also hyperactive (Maldonado-Irizarry & Kelley, 1995; Parkinson *et al.*, 1999a; Cardinal *et al.*, 2001; Cardinal & Cheung, 2005), but they do not appear to be inattentive: accuracy in tests of visuospatial attentional function is unaffected by AcbC lesions (Cole & Robbins, 1989; Christakou *et al.*, 2004). Destruction of the AcbC does not, therefore, mimic all the signs of ADHD, but these findings suggest that the behaviour of rats with AcbC damage resembles that of humans with the hyperactive–impulsive subtype of ADHD (APA, 2000).

The present results also suggest a role for the hippocampus in self-controlled choice (Cheung & Cardinal, 2005). Although the hippocampus has long been known to have a mnemonic role, the idea that the hippocampus plays a direct role in the selection of delayed rewards over immediate rewards appears novel, especially since the hippocampus does not appear to contribute to the association of actions with their outcomes over a delay. If anything, it appears to hinder this process (Cheung & Cardinal, 2005). Structural magnetic resonance imaging (MRI) studies have not shown differences in hippocampal volume between patients with ADHD and controls (Castellanos *et al.*, 1996; Filipek *et al.*, 1997), but adolescent girls with ADHD appear to have altered hippocampal glucose metabolism (Ernst *et al.*, 1997). Alterations in hippocampal function have been observed in a number of animal models of ADHD, including the coloboma mutant mouse and the neonatal rat hypoxia model (see Davids *et al.*, 2003); focal X-irradiation of the hippocampus in rats produces hippocampal granule cell ("microneuronal") hypoplasia and a syndrome of hyperactivity that also resembles ADHD (Diaz-Granados *et al.*, 1994). Ernst *et al.* (2003) found that adults with ADHD show less of a hippocampal blood flow increase than controls in a gambling game in which subjects were required to choose cards from decks that differed in the amounts and probabilities of gains and losses, akin to the Iowa gambling task of Bechara *et al.* (1994); however, choices involving reward delays were not examined.

In contrast, damage to other regions does not produce impulsive choice: for example, although the ACC, mPFC, and AcbSh have been shown to be abnormal in disorders of impulsivity (Papa *et al.*, 1996; Carey *et al.*, 1998; Ernst *et al.*, 1998; Papa *et al.*, 1998; Bush *et al.*, 1999; Rubia *et al.*, 1999; Sadile, 2000), damage to these regions does not produce impulsive choice in rats (Cardinal *et al.*, 2001; Pothuizen *et al.*, 2005). The abnormalities of structure or function observed in these regions in ADHD brains may therefore be responsible for other features of the disorder (such as inattention or motoric disinhibition) (Muir *et al.*, 1996), 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 disorders of impulsive choice may lead to more effective therapy.

5.3.3 Adolescent impulsivity

Adolescence is a time when people are prone to taking risks and seeking novel experiences. For the majority, this period is navigated safely and much useful experience is gained, but adolescence is a period of disproportionately high morbidity and mortality due to maladaptive behaviour (USA, 2001). Although adolescents are physically stronger and more resilient than children, morbidity and mortality increase 200% during adolescence, with the majority of the serious health problems related to difficulties with the control of behaviour and emotions; these problems include, for example, suicide, homicide, depression, and abuse of alcohol, nicotine, and other drugs (Dahl, 2004; Kelley *et al.*, 2004). In particular, adolescents may make choices that are rewarding in the very short term but poor in the longer term, i.e. impulsive. The adolescent Acb differs both in dopamine function and synaptic plasticity from that of the adult (see e.g. Andersen & Teicher, 2000; Philpot *et al.*, 2001; Schramm *et al.*, 2002). The hippocampus appears to

be more sensitive to ethanol, and may be more vulnerable to ethanol neurotoxicity, during adolescence (White & Swartzwelder, 2004). Similarly, development of the PFC (Giedd, 2004), inhibitory circuits in the PFC (Lewis *et al.*, 2004), and the projection from the amygdala to the PFC (Cunningham *et al.*, 2002) proceeds through this time, with PFC responsiveness also changing (Leslie *et al.*, 2004). If and how any such changes contribute to impulsive behaviour in adolescence (Adriani & Laviola, 2003) is at present unknown, though there are also demonstrable functional improvements during this time, such as in the ability to inhibit prepotent responses (see Luna & Sweeney, 2004).

5.3.4 Integration of AcbC functions with respect to impulsivity

Impulsivity is multifaceted, reflecting individual differences in distinct processes involving information gathering, the selection of outcomes, and the inhibition of motor actions (Evenden, 1999b). Furthermore, delay discounting and probability discounting may also reflect separate processes that both contribute to the selection of outcomes (see Chapter 1, p. 9). As discussed above, AcbC damage can produce impulsive choice, an impaired ability to choose delayed rewards (Cardinal *et al.*, 2001). In the context of choice involving uncertain appetitive reinforcement, "impulsivity" would equate to risk taking (less steep uncertainty discounting or greater willingness to choose unlikely rewards). AcbC lesions, however, have produced a risk-averse or conservative pattern of choice (Cardinal & Howes, 2005). Therefore, AcbC-lesioned rats cannot be characterized as impulsive in all senses. Instead, it seems that the AcbC promotes the selection, and perhaps the salience, of uncertain and delayed rewards—perhaps, in general, of rewards that are not certain, imminent, or present (Cardinal *et al.*, 2002a). The AcbC promotes choice of, and learning with, delayed rewards (Cardinal *et al.*, 2001; Cardinal & Cheung, 2005). It appears to promote the selection of uncertain reinforcers (Cardinal & Howes, 2005), and humans show increased Acb activation during the selection of high-risk options (Ernst *et al.*, 2004; Matthews *et al.*, 2004). The Acb is required for PIT, or the enhancement of instrumental responding by Pavlovian CSs signalling reward (Hall *et al.*, 2001; de Borchgrave *et al.*, 2002), for autoshaping, or approach to appetitive Pavlovian CSs (Parkinson *et al.*, 1999a; 1999b; 2000c; Cardinal *et al.*, 2002b; Parkinson *et al.*, 2002), for normal conditioned reinforcement, or working for CSs previously paired with reinforcement (Taylor & Robbins, 1984; 1986; Cador *et al.*, 1991; Parkinson *et al.*, 1999a), and Acb DA is required for the motivation to work hard (Ikemoto & Panksepp, 1999; Salamone & Correa, 2002; Salamone *et al.*, 2003; Mingote *et al.*, 2005; Salamone *et al.*, 2005).

What would one expect in an aversive context? As discussed earlier, increased probability discounting—a tendency to behave as if an uncertain outcome were less likely than it really is—would be expected to produce risk aversion for appetitive outcomes but risk proneness for aversive outcomes (Ho *et al.*, 1999). Similarly, enhanced delay discounting or temporal myopia would produce impulsive choice in an aversive context, impairing the ability to choose a small immediate penalty in preference to a large delayed penalty. In humans, at least, the delay and probability discounting processes appear similar for rewards and losses (Ostaszewski & Karzel, 2002; Green & Myerson, 2004). At present, it is not known whether AcbC lesions also affect choice involving delayed or uncertain outcomes in an aversive context; however, it is clear that the Acb is involved in aversive motivation (Salamone, 1994; Parkinson *et al.*, 1999b), including in the regulation of attention to stimuli predictive of aversive outcomes (Iordanova *et al.*, 2006).

5.3.5 Addiction

To consider the contribution of the AcbC and hippocampus to addiction, and the contribution of impulsive choice to addiction, I will first review major current theories of addiction.

5.3.5.1 Theories of addiction

In the context of the multifactorial psychological reinforcement learning framework described earlier (p. 2), the major neuropsychological theories of drug addiction—none of them mutually exclusive—can be summarized (Robbins *et al.*, 2005):

Direct positive effects of drugs; self-medication; tolerance

- Drugs are taken for their positive effects (positive reinforcement); that is, they have high instrumental incentive value. These positive effects may include euphoria, enhanced social experiences, enhanced intellectual or attentional performance, enhanced effects of other reinforcers (such as food or sex), and so on (see Wikler, 1965; 1973; Altman *et al.*, 1996; Feldman *et al.*, 1997). The precise effects depend on the drug class (Wise, 1996; Feldman *et al.*, 1997); for example, opiates such as heroin produce euphoria, and brain opioid systems may be directly involved in the assessment of "hedonic value" or pleasure (Berridge, 2000).

- An aspect of this may be that people "self-medicate" to achieve a desired level of mood, social performance, and so on (Khantzian, 1985; Weiss & Mirin, 1986; Altman *et al.*, 1996; Markou *et al.*, 1998; Newhouse *et al.*, 2004), although the extent to which self-medication of overt psychopathology occurs is debated (e.g. Castaneda *et al.*, 1994; Newhouse *et al.*, 2004). Furthermore, the effect of the drug depends upon the user's expectations (Mitchell *et al.*, 1996) and prior mood, and varies across people (Uhlenhuth *et al.*, 1981; de Wit *et al.*, 1986).

- Tolerance to pleasant drug effects may build up, requiring the user to take more drug to achieve the same effect. Tolerance can be due to a decrease in drug bioavailability ("metabolic tolerance"), a reduction in the number or responsiveness of receptors or intracellular mechanisms ("pharmacodynamic tolerance"), or a compensatory mechanism ("behavioural tolerance") (see Feldman *et al.*, 1997, p. 21). Tolerance may develop with chronic use, but in the case of cocaine, can develop in a single session (Fischman, 1989), perhaps explaining cocaine "bingeing". Metabolic tolerance is seen to barbiturates, ethanol and opiates (see Feldman *et al.*, 1997, p. 21). Pharmacodynamic tolerance is seen to a wide range of drugs including barbiturates, ethanol, opiates, amphetamine, cocaine, nicotine, and caffeine (see Feldman *et al.*, 1997, p. 21). Behavioural tolerance—conditioned tolerance—has been observed to opiates, ethanol, nicotine, benzodiazepines, and other drugs (Siegel, 1975; 1976; Krasnegor, 1978; Dafters & Anderson, 1982; Siegel, 1999). Since conditioned tolerance may be situation-specific, with the context serving as a CS, the lethality of drugs may be increased if the environment changes (Siegel, 1999).

Conditioning and sensitization

- CSs associated with the pleasant aspects of drug taking may act to promote drug taking. Drug-associated cues (including mood states, people, locations, and abuse paraphernalia) may induce some of the primary effects of drugs (Kenny *et al.*, 2003), but can also induce craving in addicts, and trigger relapse (Siegel, 1988; Tiffany & Drobes, 1990; Gawin, 1991; O'Brien *et al.*, 1998). Addicts may also work directly for drug-associated stimuli (conditioned reinforcement), leading them to the primary drug reinforcer.

- Sensitization ("inverse" or "reverse" tolerance) may also occur; this is where repeated doses of a drug enhance one or more of its effects. Prototypically, moderate, spaced doses of amphetamine enhance the subsequent locomotor response to amphetamine (Robinson & Berridge, 1993; Altman *et al.*, 1996; Kalivas *et al.*, 1998). Sensitization can exhibit environmentally specific (conditioned) properties (Post & Weiss, 1988), but sensitization regimes can also induce changes in drug pharmacodynamics (Pettit *et al.*, 1990). It has been suggested that the ability of drug-associated CSs to promote drug seeking or craving also sensitizes as a consequence of repeated drug taking (Robinson & Berridge, 1993; Wyvell & Berridge, 2001). Amphetamine sensitization also enhances the subsequent development of habits (Nelson & Killcross, 2006), discussed below.

Withdrawal and conditioned withdrawal

- Some drugs, notably the opiates and alcohol, produce powerful physical withdrawal syndromes, which are aversive. Withdrawal symptoms are improved by the drug, so the drug is taken to avoid or escape from withdrawal (negative reinforcement) (Wikler, 1965; 1973). Here, incentive learning operates for drugs of abuse just as for natural reinforcers. Just as hunger, a natural motivational state, increases the hedonic impact of foodstuffs (Berridge, 1991) and this in turn teaches the animal that it is worth working for those foodstuffs more when it is hungry (Dickinson & Balleine, 1994), opiate withdrawal reflects a "new" motivational state that the animal can perceive interoceptively, and rats have to learn that heroin has a high value in the state of opiate withdrawal (Hutcheson *et al.*, 2001a). The hedonic impact of a reinforcer may be a "common currency" for determining the value of widely different reinforcers (e.g. Cabanac, 1992).
- Environmental stimuli may become associated with withdrawal (Goldberg & Schuster, 1967; O'Brien *et al.*, 1975; 1976; 1977); CSs for withdrawal may then provoke drug taking just as withdrawal itself does (Wikler, 1965; 1973).
- Drugs such as cocaine that do not produce obvious physical withdrawal syndromes may nonetheless have unpleasant after-effects on mood (dysphoria) (Koob & Bloom, 1988; Gawin, 1991; Markou & Koob, 1991; Koob *et al.*, 1997; Knackstedt *et al.*, 2002), which may promote drug taking in the same way that physical withdrawal does. "Opponent process" theories (Solomon & Corbit, 1973; 1974; Solomon, 1980a; 1980b; Koob *et al.*, 2004) use the idea that a long-lasting anhedonic or dysphoric (i.e. unpleasant) process opposes the euphoric effects of drugs, and that with chronic use, the euphoric effects diminish and the dysphoric process comes to dominate, leading to drug taking via negative reinforcement.

Habit learning

- Drugs may activate habit-learning systems directly, so that actions that led to the drug are directly reinforced, creating powerful stimulus–response habits or "involuntary" responding, faster than with natural reinforcers (O'Brien & McLellan, 1996; Tiffany & Carter, 1998; Robbins & Everitt, 1999; Everitt *et al.*, 2001; Everitt & Wolf, 2002). A hallmark of habitual (as opposed to goal-directed) responding is that it persists even if the reinforcer's value is reduced (Dickinson, 1994). Habits are sometimes thought of as "compulsive" responding when they occur at an abnormally high level, since they do not depend on the current value of the goal. Alcohol seeking may reflect primarily habitual responding (Dickinson *et al.*, 2002), and while cocaine seeking can be goal directed (Olmstead *et al.*, 2001), under some circumstances responding for cocaine can be less susceptible to devaluation of the reinforcer (that is, more habitual) than responding for natural reinforcers (Miles *et al.*, 2003). Simi-

larly, soon after acquisition, cocaine seeking behaviour is readily suppressed by an aversive CS, whereas following prolonged experience of cocaine, this conditioned suppression is lost (Vanderschuren & Everitt, 2004). Psychostimulant sensitization also enhances subsequent habit formation (Nelson & Killcross, 2006). Craving and habits both capture something of the casual definition of addiction as "compulsive" behaviour (e.g. APA, 1994; Leshner, 1997; Koob *et al.*, 1998a).

Individual vulnerability

- People who become drug addicts may be more vulnerable than other people to one or more of these neuropsychological effects, as well as being more predisposed to try drugs of abuse in the first place.

Comparison of drug taking to alternative activities

- At a higher level of analysis, with a behavioural economic perspective, addicts weight up the benefits and costs of drug taking. They may do so rationally (Stigler & Becker, 1977; Becker & Murphy, 1988), or may exhibit decision-making flaws characteristic of humans, such as focusing inappropriately on short-term rather than long-term goals and being inconsistent in their choices (Ainslie, 1975; 1992; Herrnstein & Prelec, 1992; Heyman, 1996; Rachlin, 1997; 2000a; Ainslie, 2001).
- Drug addicts may be predisposed to act even more for short-term benefit than other people, or drugs may induce decision-making deficits in regular users (Petry *et al.*, 1998; Bickel *et al.*, 1999; Madden *et al.*, 1999; Rogers *et al.*, 1999a; Volkow *et al.*, 1999; Ainslie & Monterosso, 2003; Bickel & Johnson, 2003; Mitchell, 2003; Vuchinich & Heather, 2003); for example, as discussed earlier (p. 18), there is some evidence that self-control deficits may be a reversible consequence of cigarette dependence (Bickel & Johnson, 2003).

None of these theories, or indeed levels of explanation, is adequate on its own (Heather, 1998). For example, although heroin may be taken to alleviate withdrawal, heroin self-administration can persist in the absence of withdrawal (Bozarth & Wise, 1981; 1984), and although heroin has euphoric effects, humans will work for doses that they cannot subjectively distinguish from placebo (Lamb *et al.*, 1991). However, to focus on or seek a single theory of drug addiction is to miss the point that drugs of abuse have many effects, people take drugs for many reasons, and those reasons vary across people.

5.3.5.2 Behavioural economic approaches to addiction

To bridge the gap between neuroscientific and behavioural economic approaches to addiction, a little further elaboration of the behavioural economic approach is necessary. A direct application of traditional economics to addiction is the calculation of *elasticity of demand* for goods, such as drugs. A "good" is a commodity that, all other things being equal, agents prefer more of to less. In a barter economy, and therefore in animal experiments, the "price" of a commodity has no absolute meaning; we can speak of price only in terms of what other commodities an animal will give up to obtain the good, and that may depend on the specific commodities being traded (Friedman, 1990; Rachlin, 2003; Vuchinich & Heather, 2003). In humans, elasticity has an more general meaning, since humans use a monetary economy. Money is a single commodity that is substitutable for almost all others (is fungible), so we can calculate elasticity as the change in consumption as monetary price changes. *Own-price elasticity* measures the change in consumption of a good as its price changes. Formally,

$$\varepsilon = \frac{\text{\% small change in quantity consumed}}{\text{\% small change in price}} = \frac{\Delta q}{Q} \div \frac{\Delta p}{P} = \frac{\Delta q}{\Delta p} \cdot \frac{P}{Q} = \left(\frac{P}{Q}\right)\frac{dQ}{dP}$$
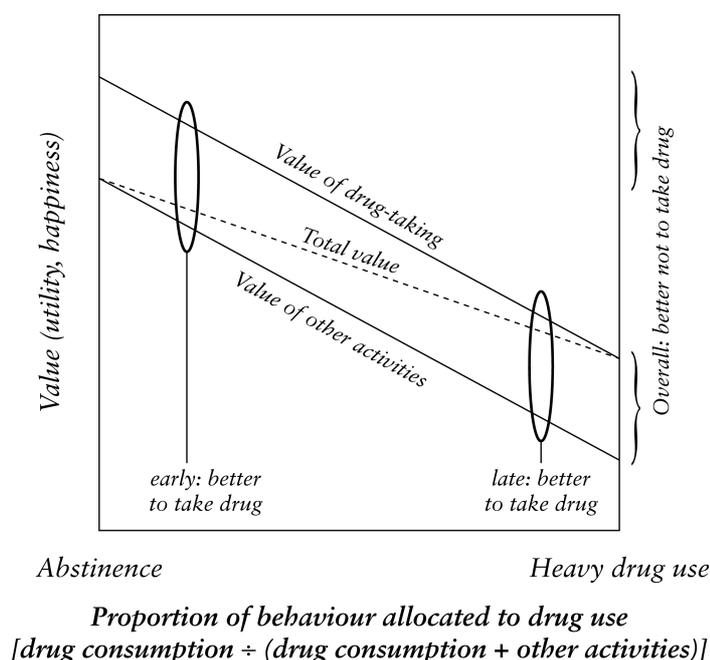
To take a simple example, suppose that biscuits cost £0.10 each, and I eat 100 biscuits/week; this costs me £10/week. If the price doubles to £0.20, I could do several things. I could halve my weekly consumption to 50 biscuits, so I continue to spend £10/week. This would be called *unit elasticity* ($\varepsilon = -1$). I could reduce my consumption by more than this, saving money. This would be *elastic* demand ($\varepsilon < -1$): my demand is very sensitive to price. Or I could not cut back very much, say to 90 biscuits/week, spending more money in total; this would be *inelastic* demand ($-1 < \varepsilon$). If demand were totally inelastic, I would consume the same no matter what the price ($\varepsilon = 0$). For most goods, demand is inelastic at low prices and elastic at high prices, termed "mixed" elasticity. *Cross-price elasticity* measures the consumption of a good as the price of other goods changes. Some commodities are *substitutes*, like butter and margarine. If the price of butter goes up, we may buy more margarine instead ($\varepsilon > 0$). Some commodities are *complements*, like gin and tonic. If the price of gin goes up, gin drinkers may buy less tonic, because they buy less gin ($\varepsilon < 0$). Some commodities are *independent*, like butter and computers, where the price of one doesn't affect consumption of the other ($\varepsilon = 0$).

An obvious way to think about addiction is that demand for drugs is inelastic compared to demand for other things. The more someone is addicted, the more inelastic their demand is; if the price increases, they will therefore sacrifice other commodities such as work, money, or social interaction, rather than sacrifice their drug. For example, alcohol demand in rats can be more inelastic than demand for food (Heyman *et al.*, 1999; Heyman, 2000). Yet drug demand is certainly not completely inelastic, and addiction is not an all-or-nothing phenomenon. Most users of heroin, cocaine, and alcohol do not use extremely large amounts, as the stereotype of an addict would suggest. Instead, most use infrequently, or "chip" (NHSDA, 2001; MacCoun, 2003b). Furthermore, most (>75%) of those dependent on an illicit drug recover (Warner *et al.*, 1995; Heyman, 2003). In fact, the elasticity of demand for cigarettes is typically about –0.4 (Gruber *et al.*, 2002; Chaloupka *et al.*, 2003); that is, if the price goes up by 10%, consumption goes down by 4%. This is for two reasons. First, when price goes up, some people quit altogether (termed *participation elasticity*). Second, people who continue to smoke, smoke less (termed *conditional elasticity of demand*, or elasticity given that someone uses the drug at all).

As for most commodities, the elasticity of drugs of abuse varies with price. Smokers working for cigarette puffs in the laboratory are fairly inelastic when the price is low ($\varepsilon = -0.56$ at a price ranging between 12–1600 responses per puff), but become more elastic when the price goes up ($\varepsilon = -1.58$ at a price ranging between 400–4500 responses per puff) (Bickel *et al.*, 1995b; DeGrandpre & Bickel, 1995; Chaloupka *et al.*, 2003). Probably for this reason, elasticity is greater for poorer smokers, for whom cigarettes are proportionally more expensive (Gruber *et al.*, 2002). In the UK, national elasticity of demand for alcohol ranges from about –1.69 for wine through –0.86 for spirits to –0.76 for beer (Smith, 1999). Participation price elasticities (the effect of price on the number of people using a drug) are about –0.90 to –0.80 for heroin and –0.55 to –0.36 for cocaine; overall elasticities (the effect of price on the total amount consumed) are about –1.80 to –1.60 for heroin and –1.10 to –0.72 for cocaine (Saffer & Chaloupka, 1995). Elasticity also varies with motivational state and other factors. Animals' demand for food is more inelastic when they are hungry and if there are no alternative ways of obtaining food (e.g. Hursh, 1978); similarly, demand for cigarettes is more inelastic when smokers have been abstinent (Madden & Bickel, 1999). When considering drug policy, it is also important to consider cross-price elasticity: if a policy reduces consumption of drug A, will the benefits be mitigated by increased consumption of drug B? In

the case of alcohol and cigarettes, the two are either complements ($\varepsilon < 0$) or independent, so reducing consumption of one tends to reduce (or not affect) consumption of the other (Gruber *et al.*, 2002). Similar analyses have been conducted for other drugs and non-drug reinforcers (Bickel *et al.*, 1995a).

Some leading economists have characterized addiction as being rational (Becker & Murphy, 1988), in that addicts take the future consequences of their behaviour into account and have stable preferences. In rational addiction theory, addiction arises because the quantities of the addictive good consumed at different time points are complements, which can lead to unstable states; this accounts, for example, for binges of consumption. Certainly, assuming rationality allows us to predict behaviour much better than not assuming rationality, unless we can predict the specific way in which people will be irrational (Friedman, 1990). A major contribution of rational addiction theory (Stigler & Becker, 1977; Becker & Murphy, 1988) was therefore to consider price as a major influence on the consumption of addictive drugs (MacCoun, 2003a). However, the premise that drug addicts choose rationally, maximizing their total happiness, has been criticized (e.g. Winston, 1980; Ainslie & Monterosso, 2003; MacCoun, 2003a). Certainly, humans do not always choose according to rational norms, as discussed in Chapter 1 (p. 1). Since hyperbolic temporal discounting is a feature of human and animal intertemporal choices (Chapter 1, p. 9), many major behavioural economic theories of addiction (Ainslie, 1975; 1992; Herrnstein & Prelec, 1992; Heyman, 1996; Rachlin, 1997; 2000a; Ainslie, 2001) emphasize that addiction results from the maximization of short-term rather than long-term utility (see MacCoun, 2003a; Vuchinich & Heather, 2003), with preferences that are inconsistent over time thanks to hyperbolic discounting, and that drug addictions


Proportion of behaviour allocated to drug use
[drug consumption ÷ (drug consumption + other activities)]

Figure 53: The "primrose path" to addiction

Good now, bad in the long run—the "primrose path" to addiction (Herrnstein & Prelec, 1992; Rachlin, 1997; 2003; Vuchinich & Heather, 2003). At any point, drug taking has a higher value than other activities, so you take the drug. But drug taking lowers both the value of future drug taking: for example, alcohol consumption causes tolerance, meaning that future alcohol isn't worth as much. Moreover, drug taking lowers the value of other activities: for example, the more alcohol you consume, the less you socialize and the worse you are at socializing; the more heroin you take, the worse you are at your job. So as you drink more, your total happiness goes down: you'd be better off not being an alcoholic. But even when you are an alcoholic, drinking now is worth more than not drinking now, for you are sensitive to local, not global, utility. As Rachlin (2000b) puts it: "The alcoholic does not choose to be an alcoholic. Instead he chooses to drink now, and now, and now, and now. The pattern of alcoholism emerges in his behaviour… without ever having been chosen."

(Rachlin, 2003) are bad ("negative" addictions) because short-term selection of drugs leads to lower long-term overall utility. Consumption of drugs reduces the value of future activities—the "primrose path" to addiction (Figure 53). Knowledge of one's own predisposition to be temporally inconsistent allows the use of self-control strategies (Ainslie, 2001; Ainslie & Monterosso, 2003; Homer, ~800 BC / 1996), such as precommitment to a particular course of action, which improve long-term utility.

Economic theories of addiction are also relevant when considering the extent to which drug use is voluntary. The diagnostic criteria for drug dependence (APA, 2000) include a compulsion to take the drug, yet drug use can certainly be voluntary. Drug use certainly has utility to the user; this may be in the form of euphoria, enhanced social experiences, or enhanced intellectual performance, depending on the drug (see Feldman *et al.*, 1997). It is debatable whether even addicts take drugs involuntarily: just because someone says they don't want to smoke and then later smokes doesn't mean they're smoking involuntarily—it might simply be that they're inconsistent (Schaler, 2000; Skog, 2003). Furthermore, not everyone who smokes wants to give up. Appreciating these differences leads to a broader classification of addiction than is conventional (Figure 54).
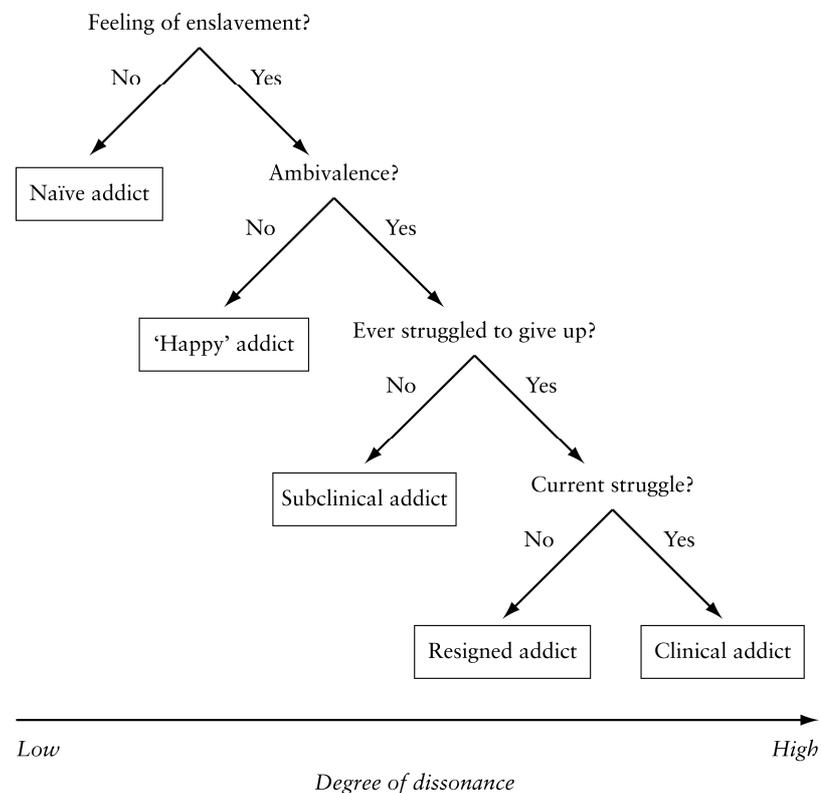
Figure 54: Skog's classification of addiction

Skog's (2003) view of addiction. A person may be unaware that it is difficult for him or her to live without a drug. Such a person is enslaved, but unaware; Skog calls them "naïve" addicts. He offers the example of a heavy drinker in Paris in World War II, who had never realised that he was dependent on alcohol until rationing came along and he was limited to one litre of wine per week. Then there are those who know that life would be harder without, but are happy with this situation: "happy" addicts, such as the 1950s smoker who thought that smoking was good for you (or at least, not bad). Those who are aware smoking is bad for you but feel no particular motivation to cut back are called "subclinical" addicts by Skog. Finally, there are those who have tried and failed but aren't trying at the moment, and those in an active struggle to quit.

The fact that people do not act to maximize their total, long-term expected reward can explain a number of otherwise counterintuitive results: for example, cigarette taxes can make smokers happier (Gruber &

Mullainathan, 2002). This implies that addiction is not "rational"—addicts' preferences are not consistent over time, and so cigarette taxes make smokers happier because they serve as a valuable self-control device, helping them to avoid smoking. Such self-control strategies are not merely a human phenomenon (Rachlin & Green, 1972; Ainslie, 1974; Ainslie & Herrnstein, 1981), as was discussed earlier. Drug addicts may discount the future more steeply (and therefore be even more impulsive and short-termist) than non-addicts (Petry *et al.*, 1998; Bickel *et al.*, 1999; Madden *et al.*, 1999; Ainslie & Monterosso, 2003; Bickel & Johnson, 2003; Mitchell, 2003; Vuchinich & Heather, 2003). Similar short-termism can explain relapse (Heyman, 2003): since one cigarette is unlikely to cause cancer and one shot of heroin doesn't condemn you to a junkie lifestyle, a person can correctly reason that since it's "just for one last time", the drug is the better choice. But a series of "one-last-times" turns into a relapse.

5.3.5.3 Contribution of the AcbC

In the context of addiction, impulsive choice plays a prominent role in maintaining the selection of drugs of abuse in favour of other, longer-term rewards (Poulos *et al.*, 1995; Heyman, 1996; Bickel *et al.*, 1999; Evenden, 1999a; Mitchell, 1999); see also p. 18. Drugs of abuse including opiates, ethanol, and psychostimulants can produce chronic adaptations in brain regions including the Acb (Koob *et al.*, 1998b), and chronic methamphetamine has been shown to increase impulsive choice in rats (Richards *et al.*, 1999a).

Furthermore, as discussed on p. 31, the motivational process provided by the AcbC (exemplified by PIT) has been suggested to be particularly significant in some addictions, perhaps becoming sensitized, and modification of this process may have therapeutic potential. If the suggestion that Pavlovian (cue-induced) motivational processes contribute to preference reversal effects and to addiction is correct (Loewenstein, 1996; Cardinal *et al.*, 2003b; Gjelsvik, 2003; Loewenstein & O'Donoghue, 2004), then the role of the AcbC is doubly important (see Cardinal *et al.*, 2002a). PIT requires the AcbC (Hall *et al.*, 2001), noncontingent CSs elevate AcbC DA levels (Bassareo & DiChiara, 1999; Ito *et al.*, 2000), DA antagonists block PIT (Dickinson *et al.*, 2000), and enhancement of Acb DA function boosts PIT (Wyvell & Berridge, 2000). PIT can also be amplified by CRH acting in the AcbSh (Pecina *et al.*, 2006), a potential mechanism through which stress may produce cue-triggered relapse in addiction. As noted above, the process of addiction is complicated further by the ability of drugs of abuse to alter the function of neural structures including the Acb (see Koob *et al.*, 1998b). 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 & 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 & Berridge, 2001).

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.

It is worth noting, however, that although a detailed knowledge of the operation of these neural systems may offer opportunities for pharmacological treatment of addiction (O'Brien, 1997), this might not change the fact that the simplest and most powerful way to influence these neural systems is often through conventional economic manipulations (MacCoun, 2003b). Nevertheless, one centre has attempted to treat addiction in humans by stereotaxic ablation of the Acb (Gao *et al.*, 2003). Citing experiments showing a reduction in heroin seeking and self-administration after AcbC lesions in rats (Alderson *et al.*, 2001; Hutcheson *et al.*, 2001b), Gao *et al.* made bilateral radiofrequency lesions of the Acb in 28 conscious recidivist opiate addicts. The authors comment that the subsequent relapse rate was markedly

lower than following the same subjects' previous detoxification attempts, after which 100% had relapsed within three weeks. Postoperatively, two subjects were lost to follow-up and were not analysed further; two (7.7% of those analysed) had relapsed within one month, 10 (38.5%) within six months, and 15 (57.7%) had relapsed by the time of publication, with those 11 (42.3%) subjects who were still abstinent beyond six months having been followed up for 8–15.5 months. Craving was apparently reduced, though no data were presented to support this assertion. This study is not a model of the scientific method: the trial was clearly neither blind nor randomized, and the control condition was the same subjects preoperatively. It is notable that the authors themselves reported improvements across their series of subjects, which they attributed to patient selection and preparation, lesion parameters, and post-discharge care; one alternative hypothesis, of course, would be that aspects of this process other than the lesion were responsible for some of the benefits. No subjects suffered intracranial bleeding or infection; side effects included "character change" (not otherwise specified) in two subjects and "slight symptoms" (not otherwise specified) in one of these, with non-disabling memory loss in four subjects. Neuropsychological data were not reported, and it is not clear whether a deficit in self control or in pursuing long-term goals was apparent.

One obvious question is raised by this set of studies. Destruction of the AcbC has been observed to reduce drug seeking in an animal model (Alderson *et al.*, 2001; Hutcheson *et al.*, 2001b); destruction of the Acb has been claimed to have similar effects in free-living humans (Gao *et al.*, 2003). Destruction of the AcbC produces impulsive choice in an animal model (Cardinal *et al.*, 2001). Impulsive choice is suggested to contribute to maladaptive behaviours including addiction (Poulos *et al.*, 1995; Heyman, 1996; Bickel *et al.*, 1999; Evenden, 1999a; Mitchell, 1999; Ainslie, 2001). Taken at face value, these three claims appear incongruent. There is insufficient evidence to resolve this question conclusively. However, the simplest explanation may be that AcbC lesions do more than produce impulsive choice. For example, such lesions impair free-operant responding for delayed rewards (Cardinal & Cheung, 2005), they impair PIT (Hall *et al.*, 2001; de Borchgrave *et al.*, 2002), and they affect conditioned reinforcement (Taylor & Robbins, 1984; 1986; Cador *et al.*, 1991; Parkinson *et al.*, 1999a), amongst other effects. In second-order schedules of reinforcement such as those used in the rat models of drug seeking cited (Alderson *et al.*, 2001; Hutcheson *et al.*, 2001b), responses are intermittently reinforced by a stimulus, and these stimuli are intermittently paired with primary reinforcement. Clearly, impairment of a cue-triggered motivational process or an inability to respond normally for conditioned reinforcers or delayed rewards might impair responding on such a schedule, independently of any effects on impulsive choice. In turn, choice impulsivity might be expected to play a more prominent role in a situation involving multiple reinforcers and genuine intertemporal choice. As was emphasized above, many factors contribute to addiction and to the selection of actions in general. It is likely that under different circumstances the AcbC both helps and hinders the pursuit of specific goals, such as a drug of abuse or any other reinforcer, as might be expected of a structure involved in making decisions about the best goal to pursue at a given moment.

5.3.5.4 Contribution of the hippocampus

Although hippocampal function is related in a number of ways to addiction, no studies to date have specifically related hippocampal dysfunction to addiction via a mechanism of impulsive choice. Likewise, while a number of studies have suggested an effect of chronic drug use on impulsivity (see p. 18), it is not clear that any such effect is mediated via the hippocampus.

The role of the hippocampus in addiction has most often been related to its role in contextual processing (see Robbins *et al.*, 2005). Theta-frequency (4–7 Hz) burst stimulation of the hippocampus (specifically, 100 Hz stimulation in five-pulse trains repeated at 5 Hz) has been shown to reinstate extinguished cocaine seeking in a manner that depended on glutamate transmission in the VTA. This has been sug-

gested to mimic the process by which reinstatement occurs when animals are placed in a context associated with drug taking, rather than in response to discrete cocaine cues (Vorel *et al.*, 2001). Dorsal hippocampal inactivation attenuates context-induced reinstatement of drug seeking, as does inactivation of the dorsal mPFC (Fuchs *et al.*, 2005). Hippocampal activity also correlates with the euphoriant effects of heroin (Sell *et al.*, 2000) and with craving for cocaine (Kilts *et al.*, 2001) or alcohol (Schneider *et al.*, 2001). Smoking-associated cues trigger hippocampal activation in nicotine-deprived smokers (Due *et al.*, 2002). Rats will self-stimulate the hippocampus electrically (Ursin *et al.*, 1966; Campbell *et al.*, 1978; Collier & Routtenberg, 1984; Campbell & Milgram, 1985) and will self-administer opioids into the hippocampus (Stevens *et al.*, 1991). Neonatal ventral hippocampal lesions have been shown to enhance simple instrumental conditioning for sucrose or cocaine subsequently in life (Chambers & Self, 2002), though the significance of this is unclear and the effects of neonatal ventral hippocampal lesions upon adult behaviour differ from the effects of lesions made in the adult (see Chambers & Self, 2002). Moreover, hippocampal, amygdala and PFC projections interact in the Acb in a way that is modulated by mesolimbic DA and that, in turn, can modulate the release of DA and influence input from other afferents to the Acb (O'Donnell & Grace, 1995; Blaha *et al.*, 1997; Floresco *et al.*, 1998; Floresco *et al.*, 2001a; Floresco *et al.*, 2001b). Thus, the hippocampus, amygdala, and PFC may influence drug seeking through their convergent projections to the Acb.

Hippocampal structure and function is altered by certain drugs of abuse, including cocaine (Thompson *et al.*, 2004; Yamaguchi *et al.*, 2004; Uz *et al.*, 2005; Yamaguchi *et al.*, 2005), nicotine (Abrous *et al.*, 2002), and opiates (Pu *et al.*, 2002). In particular, adult hippocampal neurogenesis is reduced by a number of drugs of abuse (reviewed recently by Eisch & Harburg, 2006). The relationship with addiction is unclear at present; however, parallels have been drawn with depression, in which hippocampal neurogenesis is also reduced (Kempermann & Kronenberg, 2003; Malberg & Duman, 2003; Duman, 2004). In contrast, active learning and memory formation is associated with an increase in hippocampal neurogenesis (Gould & Gross, 2002; Shors *et al.*, 2002); indeed, neurogenesis may be critical for trace conditioning (Shors, 2004), discussed earlier (pp. 100, 103). Interestingly, one action of antidepressant drugs is to increase hippocampal neurogenesis and neuronal growth (Blows, 2000; Malberg & Duman, 2003; Castren, 2004). In some cases, though not all, antidepressants are an effective therapy for drug dependence (Hughes *et al.*, 2004; Szerman *et al.*, 2005).

5.3.5.5 Perspectives on preventing and treating addiction

If neurosurgery is not a panacea, then it seems likely that conventional macroeconomic and microeconomic manipulations will remain the mainstay of the prevention and treatment of addiction. Addiction is not an all-or-nothing problem, so focusing only on prevalence (the number of people using a drug) may be inappropriate. A strategy of total harm reduction should also consider ways to reduce the average quantity used and the amount of harm per use (MacCoun, 2003b).

Many neuroscientific addiction theories focus on the way in which drugs change the brain. As Kelley & Berridge (2002) recently noted, drugs may activate the same circuits as natural rewards, perhaps in a more potent manner; they may create new states, such as the motivational state of withdrawal; and they may differentially affect the balance of processes that normally contribute to responding for natural rewards, such as habits, goal-directed actions, and cue-induced motivation. There may be other effects, too. Food makes you full and exercise makes you tired, but not all drugs will satiate you to the same extent (Heyman, 2003). Acute intoxication impairs decision making, so the decision to have the sixth pint of beer may not be made in exactly the same way as the decision to have the first. Chronic use of some drugs may alter the brain so as to impair the ability to make good choices (e.g. Rogers *et al.*, 1999a).

Some forms of brain damage may make people more likely to choose impulsively, maximizing short-term rather than long-term gain (e.g. Cardinal *et al.*, 2001). Future treatment strategies may focus on these effects, attempting to reduce drug consumption and reduce the frequency of relapse. Pharmacological strategies (Altman *et al.*, 1996; O'Brien, 1997) include drug replacement (e.g. methadone, nicotine substitution patches), antagonists to block direct drug effects (e.g. naltrexone), agents that trigger illness if the abused drug is taken (disulfiram, acamprosate), drugs that reduce craving such as DA D3 partial agonists (Pilla *et al.*, 1999) and ondansetron (Johnson *et al.*, 2002), and vaccination (Kantak, 2003). Psychological strategies include cue extinction, cognitive–behavioural therapy, and perhaps erasure of drug-associated memories (Nader, 2003). Neuroscientific advances may contribute to the diagnostic process and the matching of treatments to addicts. Techniques ranging from genetics to functional neuroimaging may become useful as a way of predicting which treatments will work best for an individual patient, and in assessing the likely efficacy of that treatment at preventing relapse before the patient is discharged. Both would be important advances.

Macroeconomic approaches take a different perspective. Once addictive behaviours are recognized to be sensitive to drug price and to the relative value of drugs and other activities, it is clear that many options currently available may be further refined. The UK pursues a policy of prohibition with regard to drugs such as heroin and cocaine, intercepting ~20% of imported drugs (Shaw, 2000) and increasing the street price. In the USA, prohibition is estimated to increase the price of cocaine by a factor of 2–4 and heroin by a factor of 6–19 (Miron, 2003). The UK spends about £1 billion per year on programmes specifically to deal with illegal drugs (UK, 2000), of which £380 million is spent on reducing drug availability and £400 million on treatment. Reducing drug availability increases price, and this reduces demand for illicit drugs (Saffer & Chaloupka, 1995); however, if demand is somewhat inelastic (elasticity $|\varepsilon| < 1$), the total amount spent on drugs increases, leading to a large criminal market (up to \$500 billion per year worldwide in 1996: Keh, 1996; Streatfeild, 2001) and health costs from contaminated drugs. Treatment of addicts is cost-effective: the benefits are in health (to the addict), reduced health costs (to the state), and reduced crime and criminal justice costs. In the USA, addiction treatment programs save about \$42,000 per treated addict per year in the costs of crime and the criminal justice system (McCollister & French, 2003), compared to about \$2,000 saved per addict treated per year in health costs. About 30% of those arrested in the UK are dependent on an illegal drug (Shaw, 2000).

In contrast, the UK policy on nicotine and alcohol is to make them legally available but heavily taxed, in order to reduce consumption, ensure that drugs are uncontaminated (by criminalizing unauthorized supply), and to produce revenue that can be spend to the benefit of addicts (e.g. treatment programmes) or society at large (e.g. health care, education). Arguably, the goal of policymakers should be to maximize the overall benefit to society (Hutcheson, 1725; Hume, 1739-1740; Mill, 1863). Whether legalization or prohibition is preferable may depend on the economics of specific drugs (Clark, 2003), but the legalization-plus-taxation option is seen as strongly preferable by many economists (see Miron & Zwiebel, 1995; Becker & Becker, 1998). In the USA, it has been estimated that legalization would lead to a ~100% increase in heroin consumption and a ~50% increase in cocaine consumption (Saffer & Chaloupka, 1995), but a net benefit of \$24 billion per year (Miron & Zwiebel, 1995; Shaw, 2000). Opiates cause about 1,000 deaths per year in England and Wales, and cocaine causes about 80 deaths (Hansard, 17 July 2002). Alcohol misuse is estimated to cost the UK perhaps £20 billion per year, of which up about £1.5 billion is spent by the NHS treating alcohol-related diseases, £6.4 billion represents lost economic productivity, and about £12 billion represents the costs of alcohol-related crime (UK, 2003); estimates of the number of deaths per year in England and Wales attributable to alcohol range from 5,000–40,000 (Hansard, 17 July

2002). In contrast, alcohol taxation generates about £11 billion (Smith, 1999). Cigarette taxes currently generate about £9.5 billion, and the NHS spends £1.5 billion treating smoking-related diseases (Parrott *et al.*, 1998), with about 120,000 deaths per year attributable to smoking in the UK (Hansard, 17 July 2002). Such taxes reduce consumption and the adverse consequences of addiction (Keeler *et al.*, 1993; Madden & Bickel, 1999; Chaloupka *et al.*, 2002; 2003). There are many macroeconomic ways to increase "price", including prohibition (reduced availability, higher financial cost, fear of criminal prosecution), restrictions on sale (availability), bans on public consumption (availability, legal sanction), taxation (financial), and stigmatizing drug users (social).

When treating individual addicts, neuroscientific strategies can also be interpreted in economic terms, allowing their comparison to other macroeconomic strategies. Pharmacological techniques can already reduce the value of specific drugs. For example, methadone treats opiate withdrawal symptoms and reduces the "high" produced by concurrently administered heroin, thus reducing the value of heroin. Heroin prescriptions (Uchtenhagen, 1997) reduce the value of contaminated, street heroin. Nicotine patches treat nicotine withdrawal, reducing the value of nicotine. Disulfiram alters ethanol metabolism temporarily so that ethanol consumption induces illness; thus, disulfiram reduces the value of alcohol. Vaccination against cocaine is being tried at the moment (Kantak, 2003); this reduces the "high" and therefore the value of cocaine. All of these can be seen as self-control tactics, and depend on the choices made by the addict: because the addict would prefer a drug-free lifestyle in the long term, he deliberately adopts a strategy (e.g. taking disulfiram) that reduces the future value of his drug. It is also possible to target the brain's motivational systems directly: thus, chemicals that reduce drug seeking in animals (e.g. Pilla *et al.*, 1999) may be another line of therapy.

Better knowledge of the risks of drug taking could also help reduce the perceived value of drugs (Heyman, 2003), and effective advertising of risk should take advantage of human reasoning biases (Slovic *et al.*, 1982), such as by vivid images of the potential unpleasant outcomes of drug use (BHF, 2004). Taken to the opposite extreme, overestimation of the risks of drug taking may also help some people avoid addiction. A personal theory that cocaine use inevitably leads to full-blown destructive addiction might not be true (Warner *et al.*, 1995; NHSDA, 2001; Heyman, 2003; MacCoun, 2003b), but this belief is a self-control device that may prevent some people taking any cocaine (Ainslie, 2001). Misinformation is clearly not a useful public health strategy, since the credibility of advisers depends upon providing accurate information, but clear and vivid statements of genuine risks are of value.

Finally, the addict pays for drugs with money and therefore forfeits other alternative commodities, and may also forfeit commodities that cannot be bought with money, such as social support. Therefore, other strategies can be used to treat addiction (Rachlin, 2003). For example, making it easier for an addict to obtain substitutes for drugs can be as effective as making it harder for the addict to obtain drugs (Green & Fisher, 2000; McCollister & French, 2003; Rachlin, 2003). Rewarding abstinence directly with money or other tangible rewards also promotes abstinence (Higgins *et al.*, 2002; Heyman, 2003). Finally, addicts can use self-control techniques like precommitment to improve their sensitivity to the long term (Ainslie, 2001).

Neuroscientific research aims to understand the neural mechanisms behind addiction, including the operation of neural systems that mediate normal reinforcement and how they are affected by drugs of abuse. In the long run, this research is likely to identify a series of molecular mechanisms that operate to promote drug taking in the addicted brain. Some will prove to be therapeutic targets, for example to reduce drug craving, and may be useful in the treatment of established addiction. Some potential therapies may be specific to the effects of drugs of abuse, but others will not be—for example, reducing strong

cravings for all reinforcers, not just abused drugs. The potential to erase drug-related memories selectively (Nader, 2003) might be of substantial benefit if it can be translated to clinical practice. Other molecular markers may indicate individual vulnerability to addiction, though it is unlikely that this information will be of much practical use, except to indicate to potential users which drugs might be relatively safe to use and which would be likely to lead to strong addictions. Furthermore, techniques may become available to predict which treatments will be best suited to an individual addict by analysing the patient's genetic makeup or neural responses. Thus the most important policy decision to be made regarding the neuroscience of addiction is how much to spend on research that may lead to treatments, and how much to spend on the treatment of addicts who seek help. However, the overall level of consumption of addictive drugs, and therefore a major component of the harm to society related to such drugs, is determined instead by macroeconomic decisions about drug regulation.

5.4 CONCLUSIONS: NEURAL SYSTEMS INVOLVED IN DELAY AND RISK ASSESSMENT

A number of limbic corticostriatal structures, together with major forebrain neuromodulatory systems, play a role in learning and choice involving delayed and probabilistic rewards. The contribution of these structures is best understood for delayed reward (Figure 55), although recent functional imaging and lesion studies have examined the neuroanatomical basis of choice involving uncertain reward.

To summarize, many structures have been implicated in the processing of delayed and/or probabilistic rewards by correlative studies, including studies of abnormalities in disorders of impulsivity such as ADHD, animal single-cell recording studies, and functional imaging studies in normal humans. Not surprisingly, these include many structures that are known to convey information concerning reward value. Impulsive choice (preference for SS over LL rewards) has been induced by lesions of the AcbC, BLA, OFC, and H; self-controlled choice has been induced by lesions of the OFC and STN. Lesions of PrL/IL and ACC do not appear to affect SS/LL reward preference; lesions of the AcbSh do not affect preference between immediate/uncertain and delayed/certain rewards. Studies examining SS/LL preference with a single pair of reinforcers cannot determine whether impulsive or self-controlled choice is due to changes in delay discounting or changes in reinforcer magnitude sensitivity. There is good evidence that changes in reinforcer magnitude sensitivity are minimal following AcbC lesions, and that AcbC damage increases delay discounting. OFC lesions appear both to enhance delay discounting and alter reinforcer magnitude sensitivity. Quantitative determinations of reinforcer magnitude sensitivity following BLA, STN, and H lesions are lacking, though there is some evidence that H lesions do not affect reward magnitude processing.

Lesions of the AcbC do not only impair choice of delayed rewards, but impair instrumental conditioning specifically when reinforcers are delayed. In contrast, although H lesions produce impulsive choice in rats, to some degree they ameliorate the deleterious effects of delays on instrumental conditioning, possibly by reducing contextual competition.

Other structures may also be 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; 1998; 2000), but so do other structures including the dorsal striatum (e.g. Schultz *et al.*, 1995), implicated in the reinforcement of stimulus–response habits (see Mishkin *et al.*,
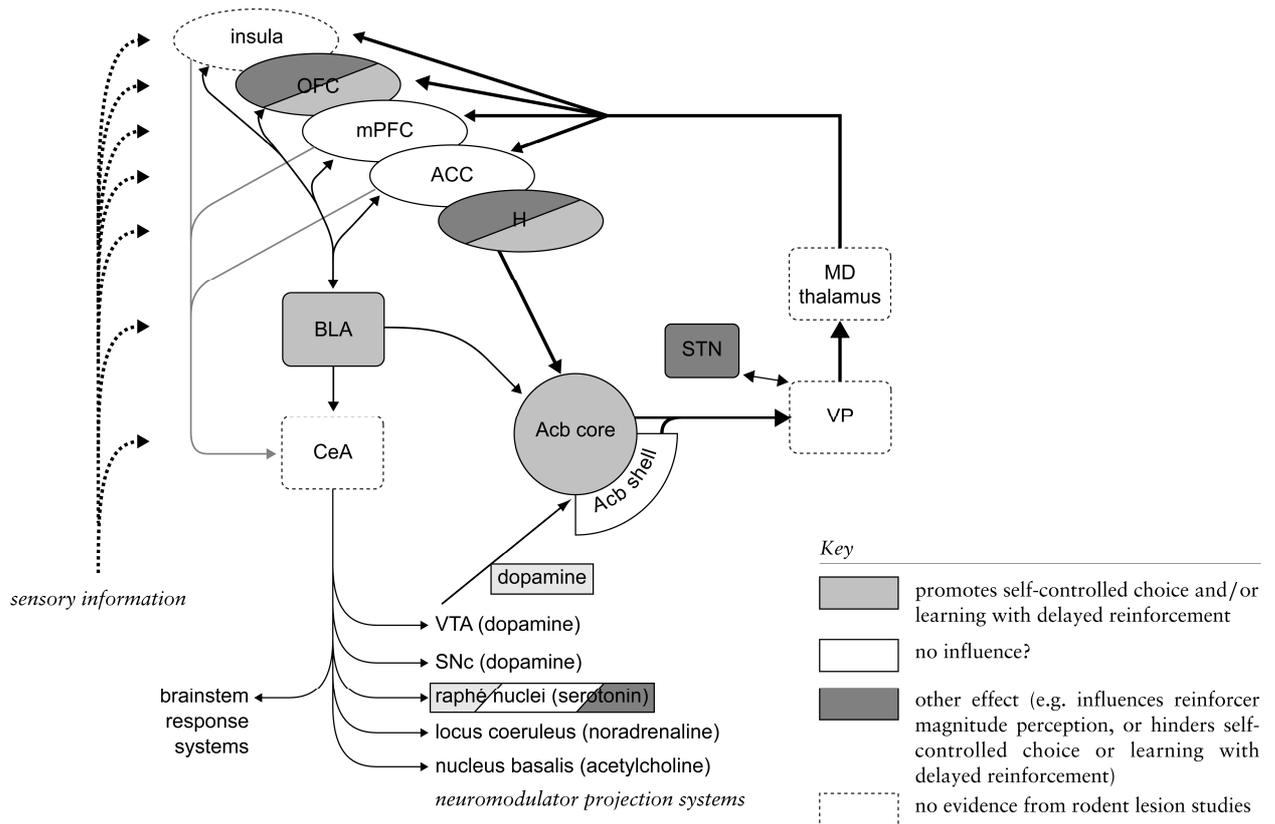
Figure 55: Key limbic corticostriatal structures involved in processing delayed reinforcement

Schematic of the limbic corticostriatal loop, showing key structures (as in Figure 7, p. 21) and their apparent influence on self-controlled choice (ability to tolerate delays to reward) as suggested by lesion studies in the rat. OFC, orbitofrontal cortex; mPFC, medial prefrontal cortex (prelimbic/infralimbic cortex in the rat); ACC, anterior cingulate cortex; H, hippocampal formation; BLA, basolateral amygdala; CeA, central nucleus of the amygdala; Acb, nucleus accumbens; STN, subthalamic nucleus; VP, ventral pallidum; MD, mediodorsal; VTA, ventral tegmental area; SNc, substantia nigra pars compacta. Not all structures and connections are shown; for example, there are projections from prefrontal cortical regions, including the OFC, to the STN (Berendse & Groenewegen, 1991; Maurice *et al.*, 1998; Hamani *et al.*, 2004).

1984; Robbins & Everitt, 1992; Packard & McGaugh, 1996; White, 1997; Parkinson *et al.*, 2000a). Furthermore, the specific pathways of communication required for choice of a delayed reward may be tested: if a structure such as the BLA or hippocampus interacts serially with the AcbC to promote choice of a delayed reward, then a disconnection lesion or inactivation (in which, for example, the BLA is inactivated in one hemisphere and the AcbC is inactivated in the other; see p. 30) should also impair subjects' ability to choose delayed reinforcement.

Neurochemically, DA D_2 receptors have been shown to promote self-controlled choice, in that D_2 antagonists have the opposite effect. NA blockade appears to affect decision making under uncertainty by reducing loss magnitude discrimination when loss probabilities are high. Forebrain 5-HT also appears to promote self-controlled choice, in that a number of studies have shown impulsive choice following 5-HT depletion or antagonists. However, not all studies have found this effect, the role of 5-HT receptor subtypes and chronic adaptations of this system is complex, and 5-HT interacts with other neuromodulators, including DA. Forebrain 5-HT depletion does not appear to alter reinforcer magnitude discrimination.

Fewer interventional studies have looked at the structures required to choose or learn from uncertain rewards, though AcbC and OFC lesions both appear to make rats less willing to choose large, uncertain rewards over small, certain rewards. 5-HT does not appear to affect choice between small, certain and

large, uncertain rewards. Human imaging studies have implicated a number of regions in decisions involving risk, including parts of the medial PFC, the Acb, and the insula. Finally, ACC lesions, BLA lesions, and ACC–BLA disconnection all appear to make rats lazy, in the sense of being less willing to choose large rewards requiring high effort to obtain, when a smaller but low-effort alternative is available.

These studies provide some insight into the pathways through which reward-related information is processed, and suggest underlying neurobiological deficits that may contribute to disorders involving risk taking and impulsive choice. Further considerations apply to drug addiction, since drugs of abuse can produce chronic adaptations in brain regions including the Acb (see Koob *et al.*, 1998b). Human addiction is associated with steep temporal discounting, particularly for the abused drug, and deficits in decision making under uncertainty. Chronic use of psychostimulants has been shown to increase impulsive choice in animal models. One mechanism contributing to addiction may therefore be the ability of drugs of abuse to induce damage or dysfunction in structures that normally promote self-controlled choice, further promoting subsequent impulsive choice and future drug taking. However, we do not yet have a mechanistic description of the way in which delays and probabilities have their effects or are encoded, or the ways in which these various limbic corticostriatal structures interact with each other to enable an animal to choose wisely.

References

Abel T, Nguyen PV, Barad M, Deuel TA, Kandel ER, Bourtchouladze R (1997). Genetic demonstration of a role for PKA in the late phase of LTP and in hippocampus-based long-term memory. *Cell* **88:** 615-626.

Abrous DN, Adriani W, Montaron MF, Aurousseau C, Rougon G, Le Moal M, Piazza PV (2002). Nicotine self-administration impairs hippocampal plasticity. *Journal of Neuroscience* **22:** 3656-3662.

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.

Adriani W, Laviola G (2003). Elevated levels of impulsivity and reduced place conditioning with d-amphetamine: two behavioral features of adolescence in mice. *Behavioral Neuroscience* **117:** 695-703.

Aggleton JP, Ed. (2000). *The amygdala: a functional analysis*. Second edition. New York: Oxford University Press.

Aggleton JP, Brown MW (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences* **22:** 425-444; discussion 444-489.

Ainslie G (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior* **21:** 485-489.

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

Ainslie G (1992). *Picoeconomics: the strategic interaction of successive motivational states within the person*. Cambridge University Press, Cambridge, UK.

Ainslie G (2001). *Breakdown of Will*. Cambridge University Press, Cambridge, UK.

Ainslie G, Herrnstein RJ (1981). Preference reversal and delayed reinforcement. *Animal Learning and Behavior* **9:** 476-482.

Ainslie G, Monterosso J (2003). Hyperbolic discounting as a factor in addiction: a critical analysis. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 35-61, 67-69. Elsevier, Oxford.

Aitken MRF, Cardinal RN (in preparation, due 2006). *Statistics for the behavioural sciences student*. Lawrence Erlbaum Associates, Mahwah, NJ.

Alderson HL, Parkinson JA, Robbins TW, Everitt BJ (2001). The effects of excitotoxic lesions of the nucleus accumbens core or shell regions on intravenous heroin self-administration in rats. *Psychopharmacology* **153:** 455-463.

Alexander GE, Crutcher MD (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences* **13:** 266-271.

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

Allison J (1979). Demand economics and experimental psychology. *Behavioral Science* **24:** 403-415.

Altman J, Everitt BJ, Glautier S, Markou A, Nutt D, Oretti R, Phillips GD, Robbins TW (1996). The biological, social and clinical bases of drug addiction: commentary and debate. *Psychopharmacology* **125:** 285-345.

Amaral DG, Witter MP (1995). Hippocampal formation. In *The Rat Nervous System*, Second edition (Paxinos G, ed.), pp. 443-493. Academic Press, London.

Amiez C, Joseph JP, Procyk E (2005a). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience* **21:** 3447-3452.

Amiez C, Joseph JP, Procyk E (2005b). Reward encoding in the monkey anterior cingulate cortex [Advance Access ahead of print; doi:10.1093/cercor/bhj046]. *Cerebral Cortex*.

Anagnostaras SG, Gale GD, Fanselow MS (2001). Hippocampus and contextual fear conditioning: recent controversies and advances. *Hippocampus* **11:** 8-17.

Anagnostaras SG, Maren S, Fanselow MS (1999). Temporally graded retrograde amnesia of contextual fear after hippocampal damage in rats: within-subjects examination. *Journal of Neuroscience* **19:** 1106-1114.

Anand BK, Brobeck JR (1951). Hypothalamic control of food intake. *Yale Journal of Biology and Medicine* **24:** 123-140.

Andersen P, Bliss TV, Lomo T, Olsen LI, Skrede KK (1969). Lamellar organization of hippocampal excitatory pathways. *Acta Physiologica Scandinavica* **76:** 4A-5A.

Andersen P, Bliss TV, Skrede KK (1971). Unit analysis of hippocampal population spikes. *Experimental Brain Research* **13:** 208-221.

Andersen SL, Teicher MH (2000). Sex differences in dopamine receptors and their relevance to ADHD. *Neuroscience and Biobehavioral Reviews* **24:** 137-141.

Anderson IM, Richell RA, Bradshaw CM (2003). The effect of acute tryptophan depletion on probabilistic choice. *Journal of Psychopharmacology* **17:** 3-7.

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

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR)*. American Psychiatric Association, Washington DC.

Apter A, Plutchik R, van Praag HM (1993). Anxiety, impulsivity and depressed mood in relation to suicidal and violent behavior. *Acta Psychiatrica Scandinavica* **87:** 1-5.

Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC (2003). Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *Journal of Neuroscience* **23:** 9632-9638.

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

Arnauld A, Nicole P (1662). *La logique, ou l'art de penser [Logic, or the Art of Thinking; the Port-Royal Logic]*. Charles Savreux, Paris.

Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology* **92:** 1144-1152.

Åsberg M (1997). Neurotransmitters and suicidal behavior. The evidence from cerebrospinal fluid studies. *Annals of the New York Academy of Sciences* **836:** 158-181.

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

Aston-Jones G, Shipley MT, Grzanna R (1995). The locus coeruleus, A5 and A7 noradrenergic cell groups. In *The Rat Nervous System* (Paxinos G, ed.), pp. 183-213. Academic Press, London.

Bäckman J, Alling C, Alsén M, Regnéll G, Träskman-Bendz L (2000). Changes of cerebrospinal fluid monoamine metabolites during long-term antidepressant treatment. *European Neuropsychopharmacology* **10:** 341-349.

Baker DA, Fuchs RA, Specio SE, Khroyan TV, Neisewander JL (1998). Effects of intraaccumbens administration of SCH-23390 on cocaine-induced locomotion and conditioned place preference. *Synapse* **30:** 181-193.

Baldwin AE, Sadeghian K, Holahan MR, Kelley AE (2002a). Appetitive instrumental learning is impaired by inhibition of cAMP-dependent protein kinase within the nucleus accumbens. *Neurobiology of Learning and Memory* **77:** 44-62.

Baldwin AE, Sadeghian K, Kelley AE (2002b). Appetitive instrumental learning requires coincident activation of NMDA and dopamine D1 receptors within the medial prefrontal cortex. *Journal of Neuroscience* **22:** 1063-1071.

Balleine B (1992). Instrumental performance following a shift in primary motivation depends on incentive learning. *Journal of Experimental Psychology: Animal Behavior Processes* **18:** 236-250.

Balleine B, Dickinson A (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *Quarterly Journal of Experimental Psychology, Section B - Comparative and Physiological Psychology* **43:** 279-296.

Balleine B, Killcross S (1994). Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. *Behavioural Brain Research* **65:** 181-193.

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

Balleine BW, Dickinson A (2000). The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. *Journal of Neuroscience* **20:** 8954-8964.

Balleine BW, Killcross AS, Dickinson A (2003). The effect of lesions of the basolateral amygdala on instrumental conditioning. *Journal of Neuroscience* **23:** 666-675.

Bannerman DM, Good MA, Butcher SP, Ramsay M, Morris RG (1995). Distinct components of spatial learning revealed by prior training and NMDA receptor blockade. *Nature* **378:** 182-186.

Bannerman DM, Yee BK, Lemaire M, Jarrard L, Iversen SD, Rawlins JN, Good MA (2001). Contextual fear conditioning is disrupted by lesions of the subcortical, but not entorhinal, connections to the hippocampus. *Experimental Brain Research* **141:** 304-311.

Barto AG (1995). Adaptive critics and the basal ganglia. In *Models of Information Processing in the Basal Ganglia* (Houk JC, Davis JL, Beiser DG, eds.), pp. 215-232. MIT Press, Cambridge, Massachusettts.

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

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

Basso AM, Kelley AE (1999). Feeding induced by GABA(A) receptor stimulation within the nucleus accumbens shell: Regional mapping and characterization of macronutrient and taste preference. *Behavioral Neuroscience* **113:** 324-336.

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.

Baunez C, Humby T, Eagle DM, Ryan LJ, Dunnett SB, Robbins TW (2001). Effects of STN lesions on simple vs choice reaction time tasks in the rat: preserved motor readiness, but impaired response selection. *European Journal of Neuroscience* **13:** 1609-1616.

Baunez C, Robbins TW (1997). Bilateral lesions of the subthalamic nucleus induce multiple deficits in an attentional task in rats. *European Journal of Neuroscience* **9:** 2086-2099.

Baxter MG, Murray EA (2001a). Effects of hippocampal lesions on delayed nonmatching-to-sample in monkeys: a reply to Zola and Squire (2001). *Hippocampus* **11:** 201-203.

Baxter MG, Murray EA (2001b). Opposite relationship of hippocampal and rhinal cortex damage to delayed nonmatching-to-sample deficits in monkeys. *Hippocampus* **11:** 61-71.

Bechara A, Damasio AR, Damasio H, Anderson SW (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50:** 7-15.

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

Bechara A, Damasio H, Tranel D, Damasio AR (1997). Deciding advantageously before knowing the advantageous strategy. *Science* **275:** 1293-1295.

Bechara A, Tranel D, Damasio H, Damasio AR (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex* **6:** 215-225.

Becker GS, Becker GM (1998). *The economics of life: From baseball to affirmative action to immigration. How real world issues affect our everyday life*. McGraw-Hill, New York.

Becker GS, Murphy KM (1988). A theory of rational addiction. *Journal of Political Economy* **96:** 675-700.

Becker SW, Brownson FO (1964). What price ambiguity? Or the role of ambiguity in decision making. *Journal of Political Economy* **72:** 62-73.

Beggs JM, Brown TH, Byrne JH, Crow T, LeDoux JE, LeBar K, Thompson RF (1999). Learning and memory: basic mechanisms. In *Fundamental Neuroscience* (Zigmond MJ, Bloom FE, Landis SC, Roberts JL, Squire LR, eds.), pp. 1411-1454. Academic Press, London.

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

Berendse HW, Galisdegraaf Y, Groenewegen HJ (1992). Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *Journal of Comparative Neurology* **316:** 314-347.

Berendse HW, Groenewegen HJ (1991). The connections of the medial part of the subthalamic nucleus in the rat: Evidence for a parallel organization. In *The Basal Ganglia III* (Bernardi G, Carpenter MB, DiChiara G, Morelli M, Stanzione P, eds.), pp. 89-98. Plenum, New York.

Berridge KC (1991). Modulation of taste affect by hunger, caloric satiety, and sensory-specific satiety in the rat. *Appetite* **16:** 103-120.

Berridge KC (2000). Measuring hedonic impact in animals and infants: microstructure of affective taste reactivity patterns. *Neuroscience and Biobehavioral Reviews* **24:** 173-198.

Berridge KC, Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews* **28:** 309-369.

Beylin AV, Gandhi CC, Wood GE, Talk AC, Matzel LD, Shors TJ (2001). The role of the hippocampus in trace conditioning: temporal discontinuity or task difficulty? *Neurobiology of Learning and Memory* **76:** 447-461.

BHF (2004). Give Up Before You Clog Up [British Heart Foundation anti-smoking advertising campaign], UK.

Bickel WK, DeGrandpre RJ, Higgins ST (1995a). The behavioral economics of concurrent drug reinforcers: a review and reanalysis of drug self-administration research. *Psychopharmacology* **118:** 250-259.

Bickel WK, DeGrandPre RJ, Higgins ST, Hughers JR, Badger G (1995b). Effects of simulated employment and recreation on drug taking: a behavioral economic analysis. *Experimental and Clinical Psychopharmacology* **3:** 467-476.

Bickel WK, Johnson MW (2003). Junk time: pathological behavior as the interaction of evolutionary and cultural forces. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 249-271, 276-278. Elsevier, Oxford.

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

Biggio G, Fadda F, Fanni P, Tagliamonte A, Gessa G (1974). Rapid depletion of serum tryptophan, brain tryptophan, serotonin and 5-hydroxyindoleacetic acid by a trytophan-free diet. *Life Sciences* **14:** 1321-1329.

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

Bjork JM, Knutson B, Fong GW, Caggiano DM, Bennett SM, Hommer DW (2004). Incentive-elicited brain activation in adolescents: similarities and differences from young adults. *Journal of Neuroscience* **24:** 1793-1802.

Blaha CD, Yang CR, Floresco SB, Barr AM, Phillips AG (1997). Stimulation of the ventral subiculum of the hippocampus evokes glutamate receptor-mediated changes in dopamine efflux in the rat nucleus accumbens. *European Journal of Neuroscience* **9:** 902-911.

Bliss TVP, Lømo T (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology* **232:** 331-356.

Blows WT (2000). The neurobiology of antidepressants. *Journal of Neuroscience Nursing* **32:** 177-180.

Boughner RL, Thomas BL, Papini MR (2004). Effects of nonreinforced preexposure to the context on autoshaping in rats: methodological implications for demonstrations of latent inhibition. *International Journal of Comparative Psychology* **17:** 168-184.

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.

Bozarth MA, Wise RA (1981). Intracranial self-administration of morphine into the ventral tegmental area in rats. *Life Sciences* **28:** 551-555.

Bozarth MA, Wise RA (1984). Anatomically distinct opiate receptor fields mediate reward and physical dependence. *Science* **224:** 516-517.

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

Bradshaw CM, 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.

Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* **30:** 619-639.

Broca P (1878). Anatomie comparée des circonvolutions cérébrales. Le grand lobe limbique et la scissure limbique dans la série des mammifères. *Revue Anthropologique (Paris)* **1:** 456-498.

Brog JS, Salyapongse A, Deutch AY, Zahm DS (1993). The patterns of afferent innervation of the core and shell in the "accumbens" part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. *Journal of Comparative Neurology* **338:** 255-278.

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

Brown PL, Jenkins HM (1968). Auto-shaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior* **11:** 1-8.

Brown VJ, 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.

Bucci DJ, Phillips RG, Burwell RD (2000). Contributions of postrhinal and perirhinal cortex to contextual information processing. *Behavioral Neuroscience* **114:** 882-894.

Buhusi CV, Meck WH (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience* **6:** 755-765.

Burk JA, Mair RG (2001). Effects of dorsal and ventral striatal lesions on delayed matching trained with retractable levers. *Behavioural Brain Research* **122:** 67-78.

Burwell RD, Saddoris MP, Bucci DJ, Wiig KA (2004). Corticohippocampal contributions to spatial and contextual learning. *Journal of Neuroscience* **24:** 3826-3836.

Bush G, Frazier JA, Rauch SL, Seidman LJ, Whalen PJ, Jenike MA, Rosen BR, Biederman J (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the Counting Stroop. *Biological Psychiatry* **45:** 1542-1552.

Bush G, Luu P, Posner MI (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences* **4:** 215-222.

Bussey TJ, Everitt BJ, 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, 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.

Cabanac M (1992). Pleasure: the common currency. *Journal of Theoretical Biology* **155:** 173-200.

Cador M, Bjijou Y, 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.

Cador M, Taylor JR, Robbins TW (1991). Potentiation of the effects of reward-related stimuli by dopaminergic-dependent mechanisms in the nucleus accumbens. *Psychopharmacology* **104:** 377-385.

Caine SB, Koob GF (1994). Effects of mesolimbic dopamine depletion on responding maintained by cocaine and food. *Journal of the Experimental Analysis of Behavior* **61:** 213-221.

Cameron AC, Windmeijer FAG (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* **77:** 329-342.

Campbell KA, Milgram NW (1985). Mechanisms underlying the plasticity of hippocampal stimulation-induced reward. *Behavioral Neuroscience* **99:** 209-219.

Campbell KA, Milgram NW, Christoff JK (1978). Plasticity in the reinforcing consequences of hippocampal stimulation. *Brain Research* **159:** 458-462.

Cardinal RN (2000). Whisker (version 1), Cambridge, UK.

Cardinal RN (2001). Neuropsychology of reinforcement processes in the rat [www.pobox.com/~rudolf/publications/2001/PhD]. Unpublished PhD thesis, University of Cambridge.

Cardinal RN (2002a). ImpulsiveChoice, computer software [www.whiskercontrol.com].

Cardinal RN (2002b). SimpleSchedules, computer software [www.whiskercontrol.com].

Cardinal RN, Aitken MRF (2001). Whisker (version 2) [www.whiskercontrol.com]. Cambridge University Technical Services Ltd, Cambridge, UK.

Cardinal RN, Aitken MRF (2006). *ANOVA for the behavioural sciences researcher*. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Cardinal RN, Cheung THC (2005). Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat. *BMC Neuroscience* **6:** 9.

Cardinal RN, Everitt BJ (2004). Neural and psychological mechanisms underlying appetitive learning: links to drug addiction. *Current Opinion in Neurobiology* **14:** 156-162.

Cardinal RN, Howes NJ (2005). Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. *BMC Neuroscience* **6:** 37.

Cardinal RN, Parkinson JA, Djafari Marbini H, Toner AJ, Bussey TJ, Robbins TW, Everitt BJ (2003a). Role of the anterior cingulate cortex in the control over behaviour by Pavlovian conditioned stimuli in rats. *Behavioral Neuroscience* **117:** 566-587.

Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002a). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews* **26:** 321-352.

Cardinal RN, Parkinson JA, Lachenal G, Halkerston KM, Rudarakanchana N, Hall J, Morrison CH, Howes SR, Robbins TW, Everitt BJ (2002b). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience* **116:** 553-567.

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

Cardinal RN, Robbins TW, 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.

Cardinal RN, Robbins TW, Everitt BJ (2003b). Choosing delayed rewards: perspectives from learning theory, neurochemistry, and neuroanatomy. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 183-213, 217-218. Elsevier, Oxford.

Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ (2004). Limbic corticostriatal systems and delayed reinforcement. *Annals of the New York Academy of Sciences* **1021:** 33-50.

Carelli RM, Wightman RM (2004). Functional microcircuitry in the accumbens underlying drug addiction: insights from real-time signaling during behavior. *Current Opinion in Neurobiology* **14:** 763-768.

Carey MP, Diewald LM, Esposito FJ, Pellicano MP, Carnevale UAG, Sergeant JA, Papa M, Sadile AG (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.

Carlson NR (1991). *Physiology of Behaviour*. Fourth edition. Allyn & Bacon, London.

Carpenter LL, Anderson GM, Pelton GH, Gudin JA, Kirwin PD, Price LH, Heninger GR, McDougle CJ (1998). Tryptophan depletion during continuous CSF sampling in healthy human subjects. *Neuropsychopharmacology* **19:** 26-35.

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301:** 386-389.

Castaneda R, Lifshutz H, Galanter M, Franco H (1994). Empirical assessment of the self-medication hypothesis among dually diagnosed inpatients. *Comprehensive Psychiatry* **35:** 180-184.

Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Vaituzis AC, Dickstein DP, Sarfatti SE, Vauss YC, Snell JW, Lange N, Kaysen D, Krain AL, Ritchie GF, Rajapakse JC, Rapoport JL (1996). Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Archives of General Psychiatry* **53:** 607-616.

Castren E (2004). Neurotrophic effects of antidepressant drugs. *Current Opinion in Pharmacology* **4:** 58-64.

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

Cervo L, Carnovali F, Stark JA, Mennini T (2003). Cocaine-seeking behavior in response to drug-associated stimuli in rats: involvement of D3 and D2 dopamine receptors. *Neuropsychopharmacology* **28:** 1150-1159.

Chaloupka FJ, Emery S, Liang L (2003). Evolving models of addictive behaviour: from neoclassical to behavioral economics. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 71-89. Elsevier, Oxford.

Chaloupka FJ, Grossman M, Saffer H (2002). The effects of price on alcohol consumption and alcohol-related problems. *Alcohol Res Health* **26:** 22-34.

Chambers RA, Self DW (2002). Motivational responses to natural and drug rewards in rats with neonatal ventral hippocampal lesions: an animal model of dual diagnosis schizophrenia. *Neuropsychopharmacology* **27:** 889-905.

Charrier D, 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, Gigerenzer G (1998). Visions of rationality. *Trends in Cognitive Sciences* **2:** 206-214.

Chen C, Kim JJ, Thompson RF, Tonegawa S (1996). Hippocampal lesions impair contextual fear conditioning in two strains of mice. *Behavioral Neuroscience* **110:** 1177-1180.

Cheung TH, Cardinal RN (2005). Hippocampal lesions facilitate instrumental learning with delayed reinforcement but induce impulsive choice in rats. *BMC Neuroscience* **6:** 36.

Childress AR, Mozley PD, McElgin W, Fitzgerald J, Reivich M, O'Brien CP (1999). Limbic activation during cue-induced cocaine craving. *American Journal of Psychiatry* **156:** 11-18.

Choi DW (1988). Glutamate neurotoxicity and diseases of the nervous system. *Neuron* **1:** 623-634.

Choi DW (1995). Calcium: still centre-stage in hypoxic-ischemic neuronal death. *Trends in Neurosciences* **18:** 58-60.

Christakou A, Robbins TW, Everitt BJ (2004). Prefrontal cortical-ventral striatal interactions involved in affective modulation of attentional performance: implications for corticostriatal circuit function. *Journal of Neuroscience* **24:** 773-780.

Christian KM, Thompson RF (2003). Neural substrates of eyeblink conditioning: acquisition and retention. *Learning & Memory* **10:** 427-455.

Clark AE (2003). The economics of drug legalization. CNRS and DELTA, France.

Clark L, Manes F, Antoun N, Sahakian BJ, Robbins TW (2003). The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia* **41:** 1474-1483.

Clark L, Roiser JP, Cools R, Rubinsztein DC, Sahakian BJ, Robbins TW (2005). Stop signal response inhibition is not modulated by tryptophan depletion or the serotonin transporter polymorphism in healthy volunteers: implications for the 5-HT theory of impulsivity. *Psychopharmacology* **182:** 570-578.

Clark RE, Manns JR, Squire LR (2001a). Trace and delay eyeblink conditioning: contrasting phenomena of declarative and nondeclarative memory. *Psychological Science* **12:** 304-308.

Clark RE, West AN, Zola SM, Squire LR (2001b). Rats with lesions of the hippocampus are impaired on the delayed nonmatching-to-sample task. *Hippocampus* **11:** 176-186.

Clemens JA, Bennett DR, Fuller RW (1980). The effect of a tryptophan-free diet on prolactin and corticosterone release by serotonergic stimuli. *Hormone and Metabolic Research* **12:** 35-38.

Coccaro EF, Siever LJ (1995). The neuropsychopharmacology of personality disorders. In *Psychopharmacology: The Fouth Generation of Progress* (Bloom FE, Kupfer DJ, eds.), Vol. 1567-1579. American College of Neuropsychopharmacology / Raven Press [http://www.acnp.org/content-32.html], New York.

Coffey SF, Gudleski GD, Saladin ME, Brady KT (2003). Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Experimental and Clinical Psychopharmacology* **11:** 18-25.

Cole BJ, 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.

Coleman-Mesches K, Salinas JA, McGaugh JL (1996). Unilateral amygdala inactivation after training attenuates memory for reduced reward. *Behavioural Brain Research* **77:** 175-180.

Collier TJ, Routtenberg A (1984). Electrical self-stimulation of dentate gyrus granule cells. *Behavioral and Neural Biology* **42:** 85-90.

Colwill RC, Rescorla RA (1986). Associative structures in instrumental learning. In *The Psychology of Learning and Motivation* (Bower GH, ed.), Vol. 20, pp. 55-104. Academic Press, Orlando, Florida.

Colwill RM, Rescorla RA (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning & Behavior* **18:** 71-82.

Cools R, Blackwell A, Clark L, Menzies L, Cox S, Robbins TW (2005). Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* **30:** 1362-1373.

Cooper SJ, Kelly CB, King DJ (1992). 5-Hydroxyindoleacetic acid in cerebrospinal fluid and prediction of suicidal behaviour in schizophrenia. *Lancet* **340:** 940-941.

Corbit LH, Balleine BW (2000). The role of the hippocampus in instrumental conditioning. *Journal of Neuroscience* **20:** 4233-4239.

Corbit LH, Balleine BW (2003). The role of prelimbic cortex in instrumental conditioning. *Behavioural Brain Research* **146:** 145-157.

Corbit LH, Muir JL, Balleine BW (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *Journal of Neuroscience* **21:** 3251-3260.

Corbit LH, Ostlund SB, Balleine BW (2002). Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *Journal of Neuroscience* **22:** 10976-10984.

Corkin S (2002). What's new with the amnesic patient H.M.? *Nature Reviews Neuroscience* **3:** 153-160.

Corkin S, Amaral DG, Gonzalez RG, Johnson KA, Hyman BT (1997). H. M.'s medial temporal lobe lesion: findings from magnetic resonance imaging. *Journal of Neuroscience* **17:** 3964-3979.

Corodimas KP, LeDoux JE (1995). Disruptive effects of posttraining perirhinal cortex lesions on conditioned fear: contributions of contextual cues. *Behavioral Neuroscience* **109:** 613-619.

Corruble E, Benyamina A, Bayle F, Falissard B, Hardy P (2003). Understanding impulsivity in severe depression? A psychometrical contribution. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **27:** 829-833.

Cousins MS, Atherton A, Turner L, Salamone JD (1996). Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behavioural Brain Research* **74:** 189-197.

Coutureau E, Dix SL, Killcross AS (2000). Involvement of the medial prefrontal cortex-basolateral amygdala pathway in fear-related behaviour in rats. *European Journal of Neuroscience* **12 (supplement 11):** 156.

Crean J, Richards JB, de Wit H (2002). Effect of tryptophan depletion on impulsive behavior in men with or without a family history of alcoholism. *Behavioural Brain Research* **136:** 349-357.

Cremniter D, Jamain S, Kollenbach K, Alvarez JC, Lecrubier Y, Gilton A, Jullien P, Lesieur P, Bonnet F, Spreux-Varoquaux O (1999). CSF 5-HIAA levels are lower in impulsive as compared to nonimpulsive violent suicide attempters and control subjects. *Biological Psychiatry* **45:** 1572-1579.

Cromwell HC, Schultz W (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *Journal of Neurophysiology* **89:** 2823-2838.

Cunningham MG, Bhattacharyya S, Benes FM (2002). Amygdalo-cortical sprouting continues into early adulthood: implications for the development of normal and abnormal function during adolescence. *Journal of Comparative Neurology* **453:** 116-130.

Dafters R, Anderson G (1982). Conditioned tolerance to the tachycardia effect of ethanol in humans. *Psychopharmacology* **78:** 365-367.

Dahl RE (2004). Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences* **1021:** 1-22.

Dalley JW, Theobald DE, Eagle DM, Passetti F, Robbins TW (2002). Deficits in impulse control associated with tonically-elevated serotonergic function in rat prefrontal cortex. *Neuropsychopharmacology* **26:** 716-728.

Daruna JH, Barnes PA (1993). A neurodevelopmental view of impulsivity. In *The impulsive client: theory, research and treatment* (McCown WG, Johnson JL, Shure MB, eds.). American Psychological Association, Washington, DC.

Datla KP, Ahier RG, Young AM, Gray JA, Joseph MH (2002). Conditioned appetitive stimulus increases extracellular dopamine in the nucleus accumbens of the rat. *European Journal of Neuroscience* **16:** 1987-1993.

Davids E, Zhang K, Tarazi FI, Baldessarini RJ (2003). Animal models of attention-deficit hyperactivity disorder. *Brain Research Reviews* **42:** 1-21.

Davis HP, Squire LR (1984). Protein synthesis and memory: a review. *Psychological Bulletin* **96:** 518-559.

Davis M (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. In *The amygdala: a functional analysis*, Second edition (Aggleton JP, ed.), pp. 213-287. Oxford University Press, New York.

Davison M, McCarthy D (1988). *The matching law: a research review*. Erlbaum, Hillsdale, NJ.

Daw ND, Niv Y, Dayan P (2005). Actions, policies, values, and the basal ganglia. In *Recent Breakthroughs in Basal Ganglia Research* (Bezard E, ed.). Nova Science Publishers, New York.

Daw ND, Touretzky DS (2002). Long-term reward prediction in TD models of the dopamine system. *Neural Computation* **14:** 2567-2583.

Day M, Langston R, Morris RG (2003). Glutamate-receptor-mediated encoding and retrieval of paired-associate learning. *Nature* **424:** 205-209.

de Borchgrave R, Rawlins JN, Dickinson A, Balleine BW (2002). Effects of cytotoxic nucleus accumbens lesions on instrumental conditioning in rats. *Experimental Brain Research* **144:** 50-68.

de la Fuente-Fernandez R, Phillips AG, Zamburlini M, Sossi V, Calne DB, Ruth TJ, Stoessl AJ (2002). Dopamine release in human ventral striatum and expectation of reward. *Behavioural Brain Research* **136:** 359-363.

de Villiers AS, Russell VA, Sagvolden T, Searson A, Jaffer A, 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, Herrnstein RJ (1976). Toward a law of response strength. *Psychological Bulletin* **83:** 1131-1153.

de Wit H, Enggasser JL, Richards JB (2002). Acute administration of d-amphetamine decreases impulsivity in healthy volunteers. *Neuropsychopharmacology* **27:** 813-825.

de Wit H, Uhlenhuth EH, Johanson CE (1986). Individual differences in the reinforcing and subjective effects of amphetamine and diazepam. *Drug and Alcohol Dependence* **16:** 341-360.

Deakin J, Aitken M, Robbins T, Sahakian BJ (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society* **10:** 590-598.

Debiec J, LeDoux JE, Nader K (2002). Cellular and systems reconsolidation in the hippocampus. *Neuron* **36:** 527-538.

DeGrandpre RJ, Bickel WK (1995). Human drug self-administration in a medium of exchange. *Experimental and Clinical Psychopharmacology* **3:** 349-357.

Delamater AR (2004). Experimental extinction in Pavlovian conditioning: behavioural and neuroscience perspectives. *Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology* **57:** 97-132.

Delay J, Brion S (1969). *Le syndrome de Korsakoff*. Masson, Paris.

Delgado PL, Charney DS, Price LH, Aghajanian GK, Landis H, Heninger GR (1990). Serotonin function and the mechanism of antidepressant action. Reversal of antidepressant-induced remission by rapid depletion of plasma tryptophan. *Archives of General Psychiatry* **47:** 411-418.

Delgado PL, Charney DS, Price LH, Landis H, Heninger GR (1989). Neuroendocrine and behavioral effects of dietary tryptophan restriction in healthy subjects. *Life Sciences* **45:** 2323-2332.

DeLong MR, Georgopoulos AP (1981). Motor functions of the basal ganglia. In *Handbook of Physiology, Section 1, The Nervous System, Vol. 2, Part 2* (Brookhart JM, Moutcastle VB, Brooks VB, eds.), pp. 1017-1061. American Physiological Society, Bethesda, Maryland.

Desmedt A, Marighetto A, Garcia R, Jaffard R (2003). The effects of ibotenic hippocampal lesions on discriminative fear conditioning to context in mice: impairment or facilitation depending on the associative value of a phasic explicit cue. *European Journal of Neuroscience* **17:** 1953-1963.

Devinsky O, Morrell MJ, Vogt BA (1995). Contributions of anterior cingulate cortex to behaviour. *Brain* **118:** 279-306.

Dhaenen H (2001). Imaging the serotonergic system in depression. *European Archives of Psychiatry and Clinical Neuroscience* **251 Suppl 2:** II76-80.

Di Chiara G (1998). A motivational learning hypothesis of the role of mesolimbic dopamine in compulsive drug use. *Journal of Psychopharmacology* **12:** 54-67.

Di Chiara G (2002). Nucleus accumbens shell and core dopamine: differential role in behavior and addiction. *Behavioural Brain Research* **137:** 75-114.

Di Ciano P, Everitt BJ (2001). Dissociable effects of antagonism of NMDA and AMPA/KA receptors in the nucleus accumbens core and shell on cocaine-seeking behavior. *Neuropsychopharmacology* **25:** 341-360.

Di Ciano P, Underwood RJ, Hagan JJ, Everitt BJ (2003). Attenuation of cue-controlled cocaine-seeking by a selective D3 dopamine receptor antagonist SB-277011-A. *Neuropsychopharmacology* **28:** 329-338.

Diaz-Granados JL, Greene PL, Amsel A (1994). Selective activity enhancement and persistence in weanling rats after hippocampal X-irradiation in infancy: possible relevance for ADHD. *Behavioral and Neural Biology* **61:** 251-259.

Dickinson A (1980). *Contemporary Animal Learning Theory*. Cambridge University Press, Cambridge.

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

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

Dickinson A, Charnock DJ (1985). Contingency effects with maintained instrumental reinforcement. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology* **37:** 397-416.

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

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

Dickinson A, Wood N, Smith JW (2002). Alcohol seeking by rats: action or habit? *Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology* **55:** 331-348.

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

Dietrich A, Allen JD, Bunnell BN (1997). Is the hippocampus involved in temporal discrimination and the memory of short intervals? *International Journal of Neuroscience* **90:** 255-269.

Draper HJ (1909). Ulysses and the Sirens.

Dudek SM, Bear MF (1992). Homosynaptic long-term depression in area CA1 of the hippocampus and the effects of NMDA receptor blockade. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 4363-4367.

Due DL, Huettel SA, Hall WG, Rubin DC (2002). Activation in mesolimbic and visuospatial neural circuits elicited by smoking cues: evidence from functional magnetic resonance imaging. *American Journal of Psychiatry* **159:** 954-960.

Duman RS (2004). Depression: a case of neuronal life and death? *Biological Psychiatry* **56:** 140-145.

Duncan CP (1949). The retroactive effect of electroconvulsive shock. *Journal of Comparative and Physiological Psychology* **42:** 32-44.

Dunnett SB (1990). Role of prefrontal cortex and striatal output systems in short-term memory deficits associated with ageing, basal forebrain lesions, and cholinergic-rich grafts. *Canadian Journal of Psychology* **44:** 210-232.

Dusek JA, Eichenbaum H (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 7109-7114.

Eagle DM, Robbins TW (2003). Lesions of the medial prefrontal cortex or nucleus accumbens core do not impair inhibitory control in rats performing a stop-signal reaction time task. *Behavioural Brain Research* **146:** 131-144.

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* **23:** 209-226.

Eisch AJ, Harburg GC (2006). Opiates, psychostimulants, and adult hippocampal neurogenesis: Insights for addiction and stem cell biology. *Hippocampus* **16:** 271-286.

Elliott R, Newman JL, Longe OA, Deakin JF (2003). Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *Journal of Neuroscience* **23:** 303-307.

Ellison GD (1964). Differential salivary conditioning to traces. *Journal of Comparative and Physiological Psychology* **57:** 373-380.

Ellsberg D (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* **75:** 643-649.

Ergorul C, Eichenbaum H (2006). Essential role of the hippocampal formation in rapid learning of higher-order sequential associations. *Journal of Neuroscience* **26:** 4111-4117.

Ernst M, Cohen RM, Liebenauer LL, Jons PH, Zametkin AJ (1997). Cerebral glucose metabolism in adolescent girls with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* **36:** 1399-1406.

Ernst M, Kimes AS, London ED, Matochik JA, Eldreth D, Tata S, Contoreggi C, Leff M, Bolla K (2003). Neural substrates of decision making in adults with attention deficit hyperactivity disorder. *American Journal of Psychiatry* **160:** 1061-1070.

Ernst M, Nelson EE, McClure EB, Monk CS, Munson S, Eshel N, Zarahn E, Leibenluft E, Zametkin A, Towbin K, Blair J, Charney D, Pine DS (2004). Choice selection and reward anticipation: an fMRI study. *Neuropsychologia* **42:** 1585-1597.

Ernst M, Zametkin AJ, Matochik JA, Jons PH, 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.

Ersche KD, Roiser JP, Clark L, London M, Robbins TW, Sahakian BJ (2005). Punishment induces risky decision-making in methadone-maintained opiate users but not in heroin users or healthy volunteers. *Neuropsychopharmacology* **30:** 2115-2124.

Erwin RJ, Ferguson ED (1979). The effect of food and water deprivation and satiation on recognition. *American Journal of Psychology* **92:** 611-626.

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.

Ettenberg A, Pettit HO, Bloom FE, Koob GF (1982). Heroin and cocaine intravenous self-administration in rats: mediation by separate neural systems. *Psychopharmacology* **78:** 204-209.

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, 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, 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.

Evenden JL, Ryan CN (1999). The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology* **146:** 413-421.

Everitt BJ, Cardinal RN, Hall J, Parkinson JA, Robbins TW (2000). Differential involvement of amygdala subsystems in appetitive conditioning and drug addiction. In *The amygdala: a functional analysis*, Second edition (Aggleton JP, ed.), pp. 353-390. Oxford University Press, New York.

Everitt BJ, Dickinson A, Robbins TW (2001). The neuropsychological basis of addictive behaviour. *Brain Research Reviews* **36:** 129-138.

Everitt BJ, Morris KA, O'Brien A, Robbins TW (1991). The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. *Neuroscience* **42:** 1-18.

Everitt BJ, Parkinson JA, Olmstead MC, Arroyo M, Robledo P, 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.

Everitt BJ, Robbins TW (2000). Second-order schedules of drug reinforcement in rats and monkeys: measurement of reinforcing efficacy and drug-seeking behaviour. *Psychopharmacology* **153:** 17-30.

Everitt BJ, Wolf ME (2002). Psychomotor stimulant addiction: a neural systems perspective. *Journal of Neuroscience* **22:** 3312-3320.

Eysenck SGB (1993). The I7: development of a measure of impulsivity and its relationship to the superfactors of personality. In *The impulsive client: theory, research and treatment* (McCown WG, Johnson JL, Shure MB, eds.). American Psychological Association, Washington DC.

Eysenck SGB, Eysenck HJ (1977). The place of impulsiveness in a dimensional system of personality description. *British Journal of Social and Clinical Psychology* **16:** 57-68.

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

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

Fellows LK, Farah MJ (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cerebral Cortex* **15:** 58-63.

Ferguson ED (2000). *Motivation: a biosocial and cognitive integration of motivation and emotion*. Oxford University Press, Oxford.

Fibiger HC, LePiane FG, Jakubovic A, Phillips AG (1987). The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *Journal of Neuroscience* **7:** 3888-3896.

Fibiger HC, Phillips AG (1988). Mesocorticolimbic dopamine systems and reward. *Annals of the New York Academy of Sciences* **537:** 206-215.

Filipek PA, Semrud-Clikeman M, Steingard RJ, Renshaw PF, Kennedy DN, Biederman J (1997). Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology* **48:** 589-601.

Fiorillo CD, Tobler PN, Schultz W (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299:** 1898-1902.

Fiorillo CD, Tobler PN, Schultz W (2005). Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behavioral and Brain Functions* **1:** 7.

Fiorino DF, Coury A, Fibiger HC, Phillips AG (1993). Electrical stimulation of reward sites in the ventral tegmental area increases dopamine transmission in the nucleus accumbens of the rat. *Behavioural Brain Research* **55:** 131-141.

Fischman MW (1989). Relationship between self-reported drug effects and their reinforcing effects: studies with stimulant drugs. *NIDA Research Monograph* **92:** 211-230.

Flexner JB, Flexner LB, Stellar E (1963). Memory in mice is affected by intracerebral puromycin. *Science* **141:** 57-58.

Flexner LB, Flexner JB, Roberts RB (1967). Memory in mice analyzed with antibiotics. Antibiotics are useful to study stages of memory and to indicate molecular events which sustain memory. *Science* **155:** 1377-1383.

Floresco SB, Blaha CD, Yang CR, Phillips AG (2001a). Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *Journal of Neuroscience* **21:** 2851-2860.

Floresco SB, Ghods-Sharifi S (2006). Amygdala-prefrontal cortical circuitry regulates effort-based decision making [Advance Access ahead of print; doi:10.1093/cercor/bhj143]. *Cerebral Cortex*.

Floresco SB, Todd CL, Grace AA (2001b). Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *Journal of Neuroscience* **21:** 4915-4922.

Floresco SB, Yang CR, Phillips AG, Blaha CD (1998). Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *European Journal of Neuroscience* **10:** 1241-1251.

Fone KC, Nutt DJ (2005). Stimulants: use and abuse in the treatment of attention deficit hyperactivity disorder. *Current Opinion in Pharmacology* **5:** 87-93.

Fortin NJ, Agster KL, Eichenbaum HB (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience* **5:** 458-462.

Fortin NJ, Wright SP, Eichenbaum H (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature* **431:** 188-191.

Friedman DD (1990). *Price Theory: An Intermediate Text [* http://www.daviddfriedman.com/Academic/Price_Theory/PThy_ToC.html *]*. South-Western Publishing Company, Cincinnati, Ohio.

Fuchs RA, Evans KA, Ledford CC, Parker MP, Case JM, Mehta RH, See RE (2005). The role of the dorsomedial prefrontal cortex, basolateral amygdala, and dorsal hippocampus in contextual reinstatement of cocaine seeking in rats. *Neuropsychopharmacology* **30:** 296-309.

Gaffan D (1992). Amnesia for complex naturalistic scenes and for objects following fornix transection in the rhesus monkey. *European Journal of Neuroscience* **4:** 381-388.

Gaffan D, Harrison S (1989). Place memory and scene memory: effects of fornix transection in the monkey. *Experimental Brain Research* **74:** 202-212.

Gallistel CR (1994). Space and time. In *Animal Learning and Cognition* (Mackintosh NJ, ed.), pp. 221-253. Academic Press, San Diego.

Gao G, Wang X, He S, Li W, Wang Q, Liang Q, Zhao Y, Hou F, Chen L, Li A (2003). Clinical study for alleviating opiate drug psychological dependence by a method of ablating the nucleus accumbens with stereotactic surgery. *Stereotactic and Functional Neurosurgery* **81:** 96-104.

Garavan H, Pankiewicz J, Bloom A, Cho JK, Sperry L, Ross TJ, Salmeron BJ, Risinger R, Kelley D, Stein EA (2000). Cue-induced cocaine craving: Neuroanatomical specificity for drug users and drug stimuli. *American Journal of Psychiatry* **157:** 1789-1798.

Garcia J (1989). Food for Tolman: Cognition and cathexis in concert. In *Aversion, avoidance and anxiety* (Archer T, Nilsson L-G, eds.), pp. 45-85. Erlbaum, Hillsdale, New Jersey.

Garcia R, Vouimba RM, Baudry M, Thompson RF (1999). The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature* **402:** 294-296.

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

George S, Rogers RD, Duka T (2005). The acute effect of alcohol on decision making in social drinkers. *Psychopharmacology* **182:** 160-169.

Gerfen CR (1992a). The neostriatal mosaic: multiple levels of compartmental organization. *Trends in Neurosciences* **15:** 133-139.

Gerfen CR (1992b). The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annual Review of Neuroscience* **15:** 285-320.

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

Giedd JN (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Sciences* **1021:** 77-85.

Giertler C, Bohn I, Hauber W (2003). The rat nucleus accumbens is involved in guiding of instrumental responses by stimuli predicting reward magnitude. *European Journal of Neuroscience* **18:** 1993-1996.

Giertler C, Bohn I, Hauber W (2004). Transient inactivation of the rat nucleus accumbens does not impair guidance of instrumental behaviour by stimuli predicting reward magnitude. *Behavioural Pharmacology* **15:** 55-63.

Giertler C, Bohn I, Hauber W (2005). Involvement of NMDA and AMPA/KA receptors in the nucleus accumbens core in instrumental learning guided by reward-predictive cues. *European Journal of Neuroscience* **21:** 1689-1702.

Gilbert PE, Kesner RP (2002). The amygdala but not the hippocampus is involved in pattern separation based on reward value. *Neurobiology of Learning and Memory* **77:** 338-353.

Gjelsvik O (2003). Reason and addiction. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 219-238, 245-247. Elsevier, Oxford.

Goldberg SR, Schuster CR (1967). Conditioned suppression by a stimulus associated with nalorphine in morphine-dependent monkeys. *Journal of the Experimental Analysis of Behavior* **10:** 235-242.

Good M (2002). Spatial memory and hippocampal function: where are we now? *Psicológica* **23:** 109-138.

Good M, Honey RC (1991). Conditioning and contextual retrieval in hippocampal rats. *Behavioral Neuroscience* **105:** 499-509.

Good M, Honey RC (1997). Dissociable effects of selective lesions to hippocampal subsystems on exploratory behavior, contextual learning, and spatial learning. *Behavioral Neuroscience* **111:** 487-493.

Gottfried JA, Dolan RJ (2004). Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. *Nature Neuroscience* **7:** 1144-1152.

Gottfried JA, O'Doherty J, Dolan RJ (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301:** 1104-1107.

Gould E, Gross CG (2002). Neurogenesis in adult mammals: some progress and problems. *Journal of Neuroscience* **22:** 619-623.

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.

Green L, Fisher EB (2000). Economic substitutability: some implications for health behavior. In *Reframing health behavior change with behavioral economics* (Bickel WK, Vuchinich RE, eds.), pp. 115-144. Erlbaum, Mahwah, NJ.

Green L, Fisher EB, Perlow S, Sherman L (1981). Preference reversal and self control: choice as a function of reward amount and delay. *Behaviour Analysis Letters* **1:** 43-51.

Green L, Myerson J (1996). Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *American Zoologist* **36:** 496-505.

Green L, Myerson J (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin* **130:** 769-792.

Green L, Myerson J, Ostaszewski P (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology. Learning, Memory, and Cognition* **25:** 418-427.

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

Griffiths DP, Dickinson A, Clayton NS (1999). Episodic memory: what can animals remember about their past? *Trends in Cognitive Sciences* **3:** 74-80.

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

Grossberg S (1982). *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control*. Reidel, Boston.

Grove EA, Domesick VB, Nauta WJ (1986). Light microscopic evidence of striatal input to intrapallidal neurons of cholinergic cell group Ch4 in the rat: a study employing the anterograde tracer Phaseolus vulgaris leucoagglutinin (PHA-L). *Brain Research* **367:** 379-384.

Gruber J, Mullainathan S (2002). Do cigarette taxes make smokers happier? [Working Paper 8872]. National Bureau of Economic Research.

Gruber J, Sen A, Stabile M (2002). Estimating price elasticities when there is smuggling: the sensitivity of smoking to price in Canada [Working Paper 8962]. National Bureau of Economic Research.

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

Hall G (1994). Pavlovian conditioning: laws of association. In *Animal Learning and Cognition* (Mackintosh NJ, ed.), pp. 15-43. Academic Press, San Diego.

Hall J, Parkinson JA, Connor TM, Dickinson A, 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, Paxinos G (1995). Serotonin and tachykinin systems. In *The Rat Nervous System* (Paxinos G, ed.), pp. 929-974. Academic Press, London.

Hamani C, Saint-Cyr JA, Fraser J, Kaplitt M, Lozano AM (2004). The subthalamic nucleus in the context of movement disorders. *Brain* **127:** 4-20.

Hansard (17 July 2002). The United Kingdom Parliament.

Harker GS (1956). Delay of reward and performance of an instrumental response. *Journal of Experimental Psychology* **51:** 303-310.

Harmer CJ, Phillips GD (1999). Enhanced dopamine efflux in the amygdala by a predictive, but not a nonpredictive, stimulus: Facilitation by prior repeated D-amphetamine. *Neuroscience* **90:** 119-130.

Hata T, Okaichi H (1998). [Effects of fimbria-fornix lesion on the temporal discrimination revealed by peak interval procedure in rats]. *Shinrigaku Kenkyu (Japanese Journal of Psychology)* **69:** 304-309.

Hauber W, Bohn I, Giertler C (2000). NMDA, but not dopamine D(2), receptors in the rat nucleus accumbens are involved in guidance of instrumental behavior by stimuli predicting reward magnitude. *Journal of Neuroscience* **20:** 6282-6288.

Haykin S (1999). *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Upper Saddle River, New Jersey.

Heather N (1998). A conceptual framework for explaining drug addiction. *Journal of Psychopharmacology* **12:** 3-7.

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

Heimer L, Zahm DS, Alheid GF (1995). Basal ganglia. In *The Rat Nervous System* (Paxinos G, ed.), pp. 579-628. Academic Press, London.

Heldt SA, Coover GD, Falls WA (2002). Posttraining but not pretraining lesions of the hippocampus interfere with feature-negative discrimination of fear-potentiated startle. *Hippocampus* **12:** 774-786.

Hellemans KG, Nobrega JN, Olmstead MC (2005). Early environmental experience alters baseline and ethanol-induced cognitive impulsivity: relationship to forebrain 5-HT1A receptor binding. *Behavioural Brain Research* **159:** 207-220.

Hernandez PJ, Sadeghian K, Kelley AE (2002). Early consolidation of instrumental learning requires protein synthesis in the nucleus accumbens. *Nature Neuroscience* **5:** 1327-1331.

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.

Herrnstein RJ, Prelec D (1992). A theory of addiction. In *Choice Over Time* (Loewenstein G, Elster J, eds.), pp. 331-361. Russell Sage Press, New York.

Hetherington AW, Ranson SW (1939). Experimental hypothalamohypophyseal obesity in the rat. *Proceedings of the Society for Experimental Biology and Medicine* **41:** 465-466.

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

Heyman GM (2000). An economic approach to animal models of alcoholism. *Alcohol Res Health* **24:** 132-139.

Heyman GM (2003). Consumption dependent changes in reward value: a framework for understanding addiction. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 95-121. Elsevier, Oxford.

Heyman GM, Gendel K, Goodman J (1999). Inelastic demand for alcohol in rats. *Psychopharmacology* **144:** 213-219.

Higgins ST, Alessi SM, Dantona RL (2002). Voucher-based incentives. A substance abuse treatment innovation. *Addictive Behaviors* **27:** 887-910.

Hikosaka K, Watanabe M (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cerebral Cortex* **10:** 263-271.

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

Hirsh R (1974). The hippocampus and contextual retrieval of information from memory: a theory. *Behavioral Biology* **12:** 421-444.

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

Holland PC, Bouton ME (1999). Hippocampus and context in classical conditioning. *Current Opinion in Neurobiology* **9:** 195-202.

Holland PC, Gallagher M (2004). Amygdala-frontal interactions and reward expectancy. *Current Opinion in Neurobiology* **14:** 148-155.

Holland PC, Gallagher M (2006). Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. *Journal of Neuroscience* **26:** 3791-3797.

Holland PC, Lamoureux JA, Han JS, Gallagher M (1999). Hippocampal lesions interfere with Pavlovian negative occasion setting. *Hippocampus* **9:** 143-157.

Holt DD, Green L, Myerson J (2003). Is discounting impulsive? Evidence from temporal and probability discounting in gambling and non-gambling college students. *Behavioural Processes* **64:** 355-367.

Holyoak KJ, Spellman BA (1993). Thinking. *Annual Review of Psychology* **44:** 265-315.

Homer (~800 BC / 1996). *Odyssey [translated by R. Fagles]*. The Bath Press, Bath.

Honey RC, Good M (1993). Selective hippocampal lesions abolish the contextual specificity of latent inhibition and conditioning. *Behavioral Neuroscience* **107:** 23-33.

Horvitz JC (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96:** 651-656.

Houk JC, Adams JL, Barto AG (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of information processing in the basal ganglia* (Houk JC, Davis JL, Beiser DG, eds.), pp. 249-270. MIT Press, Cambridge, Massachusetts / London.

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

Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310:** 1680-1683.

Hughes J, Stead L, Lancaster T (2004). Antidepressants for smoking cessation. *Cochrane Database Syst Rev***: CD000031.

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.

Hume D (1739-1740). *A Treatise of Human Nature [* http://www.gutenberg.org/etext/4705 *]*, London.

Hursh SR (1978). The economics of daily consumption controlling food- and water-reinforced responding. *Journal of the Experimental Analysis of Behavior* **29:** 475-491.

Hutcheson DM, Everitt BJ, Robbins TW, Dickinson A (2001a). The role of withdrawal in heroin addiction: enhances reward or promotes avoidance? *Nature Neuroscience* **4:** 943-947.

Hutcheson DM, Parkinson JA, Robbins TW, Everitt BJ (2001b). The effects of nucleus accumbens core and shell lesions on intravenous heroin self-administration and the acquisition of drug-seeking behaviour under a second-order schedule of heroin reinforcement. *Psychopharmacology* **153:** 464-472.

Hutcheson F (1725). *Inquiry concerning Moral Good and Evil*, London.

Huynh H, Feldt LS (1970). Conditions under which mean square ratios in repeated measures designs have exact *F*-distributions. *Journal of the American Statistical Association* **65:** 1582-1589.

Šidák Z (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* **81:** 826-831.

Ikemoto S, Panksepp J (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews* **31:** 6-41.

Iordanova MD, McNally GP, Westbrook RF (2006). Opioid receptors in the nucleus accumbens regulate attentional learning in the blocking paradigm. *Journal of Neuroscience* **26:** 4036-4045.

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

Ito R, Dalley JW, Howes SR, Robbins TW, 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.

Ito R, Dalley JW, Robbins TW, Everitt BJ (2002). Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. *Journal of Neuroscience* **22:** 6247-6253.

Ito R, Everitt BJ, Robbins TW (2005). The hippocampus and appetitive Pavlovian conditioning: effects of excitotoxic hippocampal lesions on conditioned locomotor activity and autoshaping. *Hippocampus* **15:** 713-721.

Ivkovich D, Paczkowski CM, Stanton ME (2000). Ontogeny of delay versus trace eyeblink conditioning in the rat. *Developmental Psychobiology* **36:** 148-160.

Izawa E, Aoki N, Matsushima T (2005). Neural correlates of the proximity and quantity of anticipated food rewards in the ventral striatum of domestic chicks. *European Journal of Neuroscience* **22:** 1502-1512.

Izawa E, Zachar G, Yanagihara S, Matsushima T (2003). Localized lesion of caudal part of lobus parolfactorius caused impulsive choice in the domestic chick: evolutionarily conserved function of ventral striatum. *Journal of Neuroscience* **23:** 1894-1902.

Izquierdo A, Murray EA (2005). Opposing effects of amygdala and orbital prefrontal cortex lesions on the extinction of instrumental responding in macaque monkeys. *European Journal of Neuroscience* **22:** 2341-2346.

Jaffard R, Meunier M (1993). Role of the hippocampal formation in learning and memory. *Hippocampus* **3 Spec No:** 203-217.

Jarrard LE (1993). On the role of the hippocampus in learning and memory in the rat. *Behavioral and Neural Biology* **60:** 9-26.

Jarrard LE, Meldrum BS (1993). Selective excitotoxic pathology in the rat hippocampus. *Neuropathology and Applied Neurobiology* **19:** 381-389.

Jiménez-Castellanos J, Graybiel AM (1989). Evidence that histochemically distinct zones of the primate substantia nigra pars compacta are related to patterned distributions of nigrostriatal projection neurons and striatonigral fibres. *Experimental Brain Research* **74:** 227-238.

Johansen EB, Aase H, Meyer A, Sagvolden T (2002). Attention-deficit/hyperactivity disorder (ADHD) behaviour explained by dysfunctioning reinforcement and extinction processes. *Behavioural Brain Research* **130:** 37-45.

Johnson BA, Roache JD, Ait-Daoud N, Zanca NA, Velazquez M (2002). Ondansetron reduces the craving of biologically predisposed alcoholics. *Psychopharmacology* **160:** 408-413.

Jongen-Relo AL, Kaufmann S, Feldon J (2003). A differential involvement of the shell and core subterritories of the nucleus accumbens of rats in memory processes. *Behavioral Neuroscience* **117:** 150-168.

Joseph MH, Datla K, Young AM (2003). The interpretation of the measurement of nucleus accumbens dopamine by in vivo dialysis: the kick, the craving or the cognition? *Neuroscience and Biobehavioral Reviews* **27:** 527-541.

Judge ME, Quartermain D (1982). Characteristics of retrograde amnesia following reactivation of memory in mice. *Physiology and Behavior* **28:** 585-590.

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

Kacelnik A (1997b). Normative and descriptive models of decision making: time discounting and risk sensitivity. *CIBA Foundation Symposium* **208:** 51-67; discussion 67-70.

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

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

Kamin LJ (1965). Temporal and intensity characteristics of the conditioned stimulus. In *Classical Conditioning: A Symposium* (Prokasy WF, ed.), pp. 118-147. Appleton-Century-Crofts, New York.

Kandel ER (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294:** 1030-1038.

Kantak KM (2003). Vaccines against drugs of abuse: a viable treatment option? *Drugs* **63:** 341-352.

Kaplan PS, Hearst E (1982). Bridging temporal gaps between CS and US in autoshaping: insertion of other stimuli before, during, and after CS. *Journal of Experimental Psychology: Animal Behavior Processes* **8:** 187-203.

Kawagoe R, Takikawa Y, Hikosaka O (2004). Reward-predicting activity of dopamine and caudate neurons--a possible mechanism of motivational control of saccadic eye movement. *Journal of Neurophysiology* **91:** 1013-1024.

Keeler TE, Hu TW, Barnett PG, Manning WG (1993). Taxation, regulation, and addiction: a demand function for cigarettes based on time-series evidence. *J Health Econ* **12:** 1-18.

Keh DI (1996). Drug money in a changing world: economic reform and criminal finance. United Nations Office on Drugs and Crime.

Kelley AE (1999). Neural integrative activities of nucleus accumbens subregions in relation to learning and motivation. *Psychobiology* **27:** 198-213.

Kelley AE (2004). Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neuroscience and Biobehavioral Reviews* **27:** 765-776.

Kelley AE, Bakshi VP, Haber SN, Steininger TL, Will MJ, Zhang M (2002). Opioid modulation of taste hedonics within the ventral striatum. *Physiology and Behavior* **76:** 365-377.

Kelley AE, Berridge KC (2002). The neuroscience of natural rewards: relevance to addictive drugs. *Journal of Neuroscience* **22:** 3306-3311.

Kelley AE, Schochet T, Landry CF (2004). Risk taking and novelty seeking in adolescence: introduction to part I. *Annals of the New York Academy of Sciences* **1021:** 27-32.

Kelley AE, Smith-Roe SL, Holahan MR (1997). Response-reinforcement learning is dependent on N-methyl-D-aspartate receptor activation in the nucleus accumbens core. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 12174-12179.

Kelley AE, Swanson CJ (1997). Feeding induced by blockade of AMPA and kainate receptors within the ventral striatum: a microinfusion mapping study. *Behavioural Brain Research* **89:** 107-113.

Kelley AE, Will MJ, Steininger TL, Zhang M, Haber SN (2003). Restricted daily consumption of a highly palatable food (chocolate Ensure(R)) alters striatal enkephalin gene expression. *European Journal of Neuroscience* **18:** 2592-2598.

Kemp JM, Powell TPS (1971). The connections of the striatum and globus pallidus: synthesis and speculation. *Philosophical Transactions of the Royal Society of London, Series B - Biological Sciences* **262:** 441-457.

Kempermann G, Kronenberg G (2003). Depressed new neurons--adult hippocampal neurogenesis and a cellular plasticity hypothesis of major depression. *Biological Psychiatry* **54:** 499-503.

Kenny PJ, Koob GF, Markou A (2003). Conditioned facilitation of brain reward function after repeated cocaine administration. *Behavioral Neuroscience* **117:** 1103-1107.

Keppel G (1982). *Design and analysis: a researcher's handbook.* second edition. Englewood Cliffs: Prentice-Hall, London.

Kesner RP, Williams JM (1995). Memory for magnitude of reinforcement: dissociation between the amygdala and hippocampus. *Neurobiology of Learning and Memory* **64:** 237-244.

Khantzian EJ (1985). The self-medication hypothesis of addictive disorders: focus on heroin and cocaine dependence. *American Journal of Psychiatry* **142:** 1259-1264.

Kheramin S, Body S, Herrera FM, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2005). The effect of orbital prefrontal cortex lesions on performance on a progressive ratio schedule: implications for models of inter-temporal choice. *Behavioural Brain Research* **156:** 145-152.

Kheramin S, Body S, Ho M, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2003). Role of the orbital prefrontal cortex in choice between delayed and uncertain reinforcers: a quantitative analysis. *Behavioural Processes* **64:** 239-250.

Kheramin S, Body S, Ho MY, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2004). Effects of orbital prefrontal cortex dopamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology* **175:** 206-214.

Kheramin S, Body S, Mobini S, Ho MY, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2002). Effects of quinolinic acid-induced lesions of the orbital prefrontal cortex on inter-temporal choice: a quantitative analysis. *Psychopharmacology* **165:** 9-17.

Kieres AK, Hausknecht KA, Farrar AM, Acheson A, de Wit H, Richards JB (2004). Effects of morphine and naltrexone on impulsive decision making in rats. *Psychopharmacology* **173:** 167-174.

Killcross AS, Coutureau E (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex* **13:** 400-408.

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

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

Kilts CD, Schweitzer JB, Quinn CK, Gross RE, Faber TL, Muhammad F, Ely TD, Hoffman JM, Drexler KP (2001). Neural activity related to drug craving in cocaine addiction. *Archives of General Psychiatry* **58:** 334-341.

Kim JJ, Baxter MG (2001). Multiple brain-memory systems: the whole does not equal the sum of its parts. *Trends in Neurosciences* **24:** 324-330.

Kim JJ, Fanselow MS (1992). Modality-specific retrograde amnesia of fear. *Science* **256:** 675-677.

Kim JJ, Rison RA, Fanselow MS (1993). Effects of amygdala, hippocampus, and periaqueductal gray lesions on short- and long-term contextual fear. *Behavioral Neuroscience* **107:** 1093-1098.

Kirby KN, Marakovic NN (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*.

Kirby KN, Petry NM (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction* **99:** 461-471.

Kirby KN, Petry NM, Bickel WK (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General* **128:** 78-87.

Knackstedt LA, Samimi MM, Ettenberg A (2002). Evidence for opponent-process actions of intravenous cocaine and cocaethylene. *Pharmacology, Biochemistry and Behavior* **72:** 931-936.

Knutson B, Adams CM, Fong GW, Hommer D (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience* **21:** RC159.

Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005). Distributed neural representation of expected value. *Journal of Neuroscience* **25:** 4806-4812.

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

Koob GF, Ahmed SH, Boutrel B, Chen SA, Kenny PJ, Markou A, O'Dell LE, Parsons LH, Sanna PP (2004). Neurobiological mechanisms in the transition from drug use to drug dependence. *Neuroscience and Biobehavioral Reviews* **27:** 739-749.

Koob GF, Bloom FE (1988). Cellular and molecular mechanisms of drug dependence. *Science* **242:** 715-723.

Koob GF, Caine SB, Parsons L, Markou A, Weiss F (1997). Opponent process model and psychostimulant addiction. *Pharmacology Biochemistry and Behavior* **57:** 513-521.

Koob GF, Rocio M, Carrera A, Gold LH, Heyser CJ, Maldonado-Irizarry C, Markou A, Parsons LH, Roberts AJ, Schulteis G, Stinus L, Walker JR, Weissenborn R, Weiss F (1998a). Substance dependence as a compulsive behavior. *Journal of Psychopharmacology* **12:** 39-48.

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

Krasnegor NA (1978). Behavorial tolerance: research and treatment implications: introduction. *NIDA Research Monograph***:** 1-3.

Kuhnen CM, Knutson B (2005). The neural basis of financial risk taking. *Neuron* **47:** 763-770.

Kuntsi J, Oosterlaan J, Stevenson J (2001). Psychological mechanisms in hyperactivity: I. Response inhibition deficit, working memory impairment, delay aversion, or something else? *Journal of Child Psychology and Psychiatry and Allied Disciplines* **42:** 199-210.

Lagorio CH, Madden GJ (2005). Delay discounting of real and hypothetical rewards III: steady-state assessments, forced-choice trials, and all real rewards. *Behavioural Processes* **69:** 173-187.

Lamb RJ, Preston KL, Schindler CW, Meisch RA, Davis F, Katz JL, Henningfield JE, Goldberg SR (1991). The reinforcing and subjective effects of morphine in post-addicts: a dose-response study. *Journal of Pharmacology and Experimental Therapeutics* **259:** 1165-1173.

Lane SD, Cherek DR, Pietras CJ, Tcheremissine OV (2004). Alcohol effects on human risk taking. *Psychopharmacology* **172:** 68-77.

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

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

LeDoux JE (2000). The amygdala and emotion: a view through fear. In *The amygdala: a functional analysis*, Second edition (Aggleton JP, ed.), pp. 289-310. Oxford University Press, New York.

Lee HJ, Groshek F, Petrovich GD, Cantalini JP, Gallagher M, Holland PC (2005a). Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. *Journal of Neuroscience* **25:** 3881-3888.

Lee HJ, Youn JM, O MJ, Gallagher M, Holland PC (2005b). Role of amygdala-nigral connections in enhanced attention for conditioned stimulus processing. *Society for Neuroscience Abstracts* **35: 411.15**.

Lee JL, Everitt BJ, Thomas KL (2004). Independent cellular processes for hippocampal memory consolidation and reconsolidation. *Science* **304:** 839-843.

Leland DS, Paulus MP (2005). Increased risk-taking decision-making but not altered response to punishment in stimulant-using young adults. *Drug and Alcohol Dependence* **78:** 83-90.

Leon MI, Gallistel CR (1998). Self-stimulating rats combine subjective reward magnitude and subjective reward rate multiplicatively. *Journal of Experimental Psychology: Animal Behavior Processes* **24:** 265-277.

Leshner AI (1997). Addiction is a brain disease, and it matters. *Science* **278:** 45-47.

Leslie FM, Loughlin SE, Wang R, Perez L, Lotfipour S, Belluzzia JD (2004). Adolescent development of forebrain stimulant responsiveness: insights from animal studies. *Annals of the New York Academy of Sciences* **1021:** 148-159.

Levene H (1960). Robust tests for the equality of variance. In *Contributions to probability and statistics* (Oklin I, ed.). Stanford University Press, Palo Alto, California.

Levita L, Dalley JW, Robbins TW (2002). Disruption of Pavlovian contextual conditioning by excitotoxic lesions of the nucleus accumbens core. *Behavioral Neuroscience* **116:** 539-552.

Lewis DA, Cruz D, Eggan S, Erickson S (2004). Postnatal development of prefrontal inhibitory circuits and the pathophysiology of cognitive dysfunction in schizophrenia. *Annals of the New York Academy of Sciences* **1021:** 64-76.

Liao RM, Chuang FJ (2003). Differential effects of diazepam infused into the amygdala and hippocampus on negative contrast. *Pharmacology, Biochemistry and Behavior* **74:** 953-960.

Lindgren JL, Gallagher M, Holland PC (2003). Lesions of basolateral amygdala impair extinction of CS motivational value, but not of explicit conditioned responses, in Pavlovian appetitive second-order conditioning. *European Journal of Neuroscience* **17:** 160-166.

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

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

Liu YP, Wilkinson LS, Robbins TW (2004). Effects of acute and chronic buspirone on impulsive choice and efflux of 5-HT and dopamine in hippocampus, nucleus accumbens and prefrontal cortex. *Psychopharmacology* **173:** 175-185.

Loewenstein G (1996). Out of control: visceral influences on behavior. *Organizational Behavior and Human Decision Processes* **63:** 272-292.

Loewenstein GF, O'Donoghue T (2004). Animal Spirits: Affective and Deliberative Processes in Economic Behavior [http://ssrn.com/abstract=539843].

Logue AW, Tobin H, Chelonis JJ, Wang RY, Geary N, Schachter S (1992). Cocaine decreases self-control in rats: a preliminary report. *Psychopharmacology* **109:** 245-247.

Lopes LL (1994). Psychology and economics: perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology* **45:** 197-227.

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

Luna B, Sweeney JA (2004). The emergence of collaborative brain function: FMRI studies of the development of response inhibition. *Annals of the New York Academy of Sciences* **1021:** 296-309.

Maas LC, Lukas SE, Kaufman MJ, Weiss RD, Daniels SL, Rogers VW, Kukes TJ, Renshaw PF (1998). Functional magnetic resonance imaging of human brain activation during cue-induced cocaine craving. *American Journal of Psychiatry* **155:** 124-126.

MacCoun R (2003a). Comments on Chaloupka, Emery, and Liang. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 90-94. Elsevier, Oxford.

MacCoun R (2003b). Is the addiction concept useful for drug policy? In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 383-401, 407. Elsevier, Oxford.

MacCrimmon KR (1968). Descriptive and normative implications of the decision-theory postulates. In *Risk and Uncertainty* (Borch K, Mossin J, eds.), pp. 3-23. MacMillan, London.

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

Mackintosh NJ (1983). *Conditioning and associative learning*. Oxford University Press, Oxford.

MacLean PD (1949). Psychosomatic disease and the "visceral brain": recent developments bearing on the Papez theory of emotion. *Psychosomatic Medicine* **11:** 338-353.

MacLean PD (1952). Some psychiatric implications of physiological studies on the frontotemporal portion of the limbic system (visceral brain). *Electroencephalography and Clinical Neurophysiology* **4:** 407-418.

MacLean PD (1993). Introduction: perspectives on cingulate cortex in the limbic system. In *Neurobiology of cingulate cortex and limbic thalamus: a comprehensive handbook* (Vogt BA, Gabriel M, eds.), pp. 1-15. Birkhauser, Boston, Massachusetts.

Mactutus CF, Ferek JM, George CA, Riccio DC (1982). Hypothermia-induced amnesia for newly acquired and old reactivated memories: commonalities and distinctions. *Physiological Psychology* **10:** 79-95.

Madden GJ, Bickel WK (1999). Abstinence and price effects on demand for cigarettes: a behavioral-economic analysis. *Addiction* **94:** 577-588.

Madden GJ, Bickel WK, Jacobs EA (1999). Discounting of delayed rewards in opioid-dependent outpatients: exponential or hyperbolic discounting functions? *Experimental and Clinical Psychopharmacology* **7:** 284-293.

Madden GJ, Petry NM, Badger GJ, Bickel WK (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: drug and monetary rewards. *Experimental and Clinical Psychopharmacology* **5:** 256-262.

Maguire EA, Burgess N, Donnett JG, Frackowiak RS, Frith CD, O'Keefe J (1998). Knowing where and getting there: a human navigation network. *Science* **280:** 921-924.

Maguire EA, Frackowiak RS, Frith CD (1997). Recalling routes around London: activation of the right hippocampus in taxi drivers. *Journal of Neuroscience* **17:** 7103-7110.

Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, Frith CD (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 4398-4403.

Malberg JE, Duman RS (2003). Cell proliferation in adult hippocampus is decreased by inescapable stress: reversal by fluoxetine treatment. *Neuropsychopharmacology* **28:** 1562-1571.

Maldonado-Irizarry CS, Kelley AE (1995). Excitotoxic lesions of the core and shell subregions of the nucleus accumbens differentially disrupt body-weight regulation and motor activity in the rat. *Brain Research Bulletin* **38:** 551-559.

Manes F, Sahakian B, Clark L, Rogers R, Antoun N, Aitken M, Robbins T (2002). Decision-making processes following damage to the prefrontal cortex. *Brain* **125:** 624-639.

Mann JJ (2003). Neurobiology of suicidal behaviour. *Nature Reviews Neuroscience* **4:** 819-828.

Maren S (1999). Neurotoxic or electrolytic lesions of the ventral subiculum produce deficits in the acquisition and expression of Pavlovian fear conditioning in rats. *Behavioral Neuroscience* **113:** 283-290.

Maren S, Fanselow MS (1997). Electrolytic lesions of the fimbria/fornix, dorsal hippocampus, or entorhinal cortex produce anterograde deficits in contextual fear conditioning in rats. *Neurobiology of Learning and Memory* **67:** 142-149.

Markou A, Koob GF (1991). Postcocaine anhedonia. An animal model of cocaine withdrawal. *Neuropsychopharmacology* **4:** 17-26.

Markou A, Kosten TR, Koob GF (1998). Neurobiological similarities in depression and drug dependence: a self-medication hypothesis. *Neuropsychopharmacology* **18:** 135-174.

Marlin NA (1980). Contextual associations in trace conditioning. *Animal Learning and Behavior* **9:** 519-523.

Marlin NA, Miller RR (1981). Associations to contextual stimuli as a determinant of long-term habituation. *Journal of Experimental Psychology: Animal Behavior Processes* **7:** 313-333.

Marshall JF, Teitelbaum P (1977). New considerations in the neuropsychology of motivated behaviors. In *Handbook of Psychopharmacology* (Iversen LL, Iversen SD, Snyder SH, eds.), Vol. 7, pp. 201-222. Plenum Press, New York.

Martin JH (1989). *Neuroanatomy: Text and Atlas*. Appleton & Lange, East Norwalk, CT.

Martin PD, Ono T (2000). Effects of reward anticipation, reward presentation, and spatial parameters on the firing of single neurons recorded in the subiculum and nucleus accumbens of freely moving rats. *Behavioural Brain Research* **116:** 23-38.

Martin-Iverson MT, Wilkie D, Fibiger HC (1987). Effects of haloperidol and d-amphetamine on perceived quantity of food and tones. *Psychopharmacology* **93:** 374-381.

Matthews SC, Simmons AN, Lane SD, Paulus MP (2004). Selective activation of the nucleus accumbens during risk-taking decision making. *Neuroreport* **15:** 2123-2127.

Mauchly JW (1940). Significance test for sphericity of a normal *n*-variate distribution. *Annals of Mathematical Statistics* **11:** 204-209.

Maurice N, Deniau JM, Glowinski J, Thierry AM (1998). Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the corticosubthalamic circuits. *Journal of Neuroscience* **18:** 9539-9546.

Maviel T, Durkin TP, Menzaghi F, Bontempi B (2004). Sites of neocortical reorganization critical for remote spatial memory. *Science* **305:** 96-99.

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

Mazur JE (1988). Choice between small certain and large uncertain reinforcers. *Animal Learning & Behavior* **16:** 199-205.

Mazur JE (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior* **51:** 87-99.

Mazur JE (1995). Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *Journal of the Experimental Analysis of Behavior* **63:** 139-150.

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

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

McClelland JL, McNaughton BL, O'Reilly RC (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* **102:** 419-457.

McCloskey M, Cohen N, Eds. (1989). *Catastrophic interference in connectionist networks: The sequential learning problem*. The Psychology of Learning and Motivation (Vol. 24). Edited by Bower GH. New York: Academic Press.

McClure SM, Berns GS, Montague PR (2003a). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38:** 339-346.

McClure SM, Daw ND, Montague PR (2003b). A computational substrate for incentive salience. *Trends in Neurosciences* **26:** 423-428.

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* **306:** 503-507.

McCollister KE, French MT (2003). The relative contribution of outcome domains in the total economic benefit of addiction interventions: a review of first findings. *Addiction* **98:** 1647-1659.

McEchron MD, Bouwmeester H, Tseng W, Weiss C, Disterhoft JF (1998). Hippocampectomy disrupts auditory trace fear conditioning and contextual fear conditioning in the rat. *Hippocampus* **8:** 638-646.

McEchron MD, Disterhoft JF (1999). Hippocampal encoding of non-spatial trace conditioning. *Hippocampus* **9:** 385-396.

McNish KA, Gewirtz JC, Davis M (1997). Evidence of contextual fear after lesions of the hippocampus: a disruption of freezing but not fear-potentiated startle. *Journal of Neuroscience* **17:** 9353-9360.

Meck WH (1988). Hippocampal function is required for feedback control of an internal clock's criterion. *Behavioral Neuroscience* **102:** 54-60.

Meck WH, Church RM, Olton DS (1984). Hippocampus, time, and memory. *Behavioral Neuroscience* **98:** 3-22.

Mehlman PT, Higley JD, Faucher I, Lilly AA, Taub DM, Vickers J, Suomi SJ, Linnoila M (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.

Miles FJ, Everitt BJ, Dickinson A (2003). Oral cocaine seeking by rats: Action or habit? *Behavioral Neuroscience* **117:** 927-938.

Mill JS (1863). *Utilitarianism [* http://www.gutenberg.org/etext/11224 *]*, London.

Miller RR, Ott CA, Berk AM, Springer AD (1974). Appetitive memory restoration after electro-convulsive shock in the rat. *Journal of Comparative and Physiological Psychology* **87:** 717-723.

Miller RR, Springer AD (1972). Induced recovery of memory in rats following electroconvulsive shock. *Physiology and Behavior* **8:** 645-651.

Millin PM, Moody EW, Riccio DC (2001). Interpretations of retrograde amnesia: old problems redux. *Nature Reviews Neuroscience* **2:** 68-70.

Mingote S, Weber SM, Ishiwari K, Correa M, Salamone JD (2005). Ratio and time requirements on operant schedules: effort-related effects of nucleus accumbens dopamine depletions. *European Journal of Neuroscience* **21:** 1749-1757.

Minsky ML (1961). Steps towards artificial intelligence. *Proceedings of the Institute of Radio Engineers* **9:** 8-30.

Miron JA (2003). The effect of drug prohibition on drug prices: evidence from the markets for cocaine and heroin. *The Review of Economics and Statistics* **85:** 522-530.

Miron JA, Zwiebel J (1995). The economic case against drug prohibition. *Journal of Economic Perspectives* **9:** 175-192.

Misanin JR, Miller RR, Lewis DJ (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science* **160:** 554-555.

Mischel W (1966). Theory and research on the antecedents of self-imposed delay of reward. In *Progress in experimental personality research* (Maher BA, ed.), Vol. 3, pp. 85-132. Academic Press, New York.

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

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

Mitchell SH (2003). Discounting the value of commodities according to different types of cost. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 339-357. Elsevier, Oxford.

Mitchell SH (2004a). Effects of short-term nicotine deprivation on decision-making: delay, uncertainty and effort discounting. *Nicotine & Tobacco Research* **6:** 819-828.

Mitchell SH (2004b). Measuring impulsivity and modeling its association with cigarette smoking. *Behavioral and Cognitive Neuroscience Reviews* **3:** 261-275.

Mitchell SH, Laurent CL, de Wit H (1996). Interaction of expectancy and the pharmacological effects of d-amphetamine: subjective effects and self-administration. *Psychopharmacology* **125:** 371-378.

Miyazaki K, Mogi E, Araki N, Matsumoto G (1998). Reward-quality dependent anticipation in rat nucleus accumbens. *Neuroreport* **9:** 3943-3948.

Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **160:** 290-298.

Mobini S, Chiang TJ, Al-Ruwaitea AS, Ho MY, Bradshaw CM, Szabadi E (2000a). Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology* **149:** 313-318.

Mobini S, Chiang TJ, Ho MY, Bradshaw CM, Szabadi E (2000b). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **152:** 390-397.

Mogenson GJ, Jones DL, Yim CY (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in Neurobiology* **14:** 69-97.

Morgan MA, LeDoux JE (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behavioral Neuroscience* **109:** 681-688.

Morgan MA, LeDoux JE (1999). Contribution of ventrolateral prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Neurobiology of Learning and Memory* **72:** 244-251.

Morgan MA, Romanski LM, LeDoux JE (1993). Extinction of emotional learning: contribution of medial prefrontal cortex. *Neuroscience Letters* **163:** 109-113.

Morris RG (2001). Episodic-like memory in animals: psychological criteria, neural mechanisms and the value of episodic-like tasks to investigate animal models of neurodegenerative disease. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **356:** 1453-1465.

Morris RG, Anderson E, Lynch GS, Baudry M (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* **319:** 774-776.

Morris RG, Frey U (1997). Hippocampal synaptic plasticity: role in spatial learning or the automatic recording of attended experience? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **352:** 1489-1503.

Morris RG, Garrud P, Rawlins JN, O'Keefe J (1982). Place navigation impaired in rats with hippocampal lesions. *Nature* **297:** 681-683.

Morris RGM (1994). The neural basis of learning with particular reference to the role of synaptic plasticity. Where are we a century after Cajal's speculations? In *Animal Learning and Cognition* (Mackintosh NJ, ed.), pp. 135-183. Academic Press, San Diego.

Moser MB, Moser EI, Forrest E, Andersen P, Morris RG (1995). Spatial learning with a minislab in the dorsal hippocampus. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 9697-9701.

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

Muir JL, Everitt BJ, 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.

Mullainathan S (2002). Behavioral economics. In *International Encyclopedia of the Social & Behavioral Sciences* (Baltes PB, Smelser NJ, eds.). Pergamon, Oxford.

Murray EA, Bussey TJ (2001). Consolidation and the medial temporal lobe revisited: methodological considerations. *Hippocampus* **11:** 1-7.

Murray EA, Mishkin M (1998). Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *Journal of Neuroscience* **18:** 6568-6582.

Myers KM, Davis M (2002). Behavioral and neural analysis of extinction. *Neuron* **36:** 567-584.

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

Myerson J, Green L, Hanson JS, Hold DD, Estle SJ (2003). Discounting delayed and probabilistic rewards: Processes and traits. *Journal of Economic Psychology* **24:** 619-635.

Nadel L, Bohbot V (2001). Consolidation of memory. *Hippocampus* **11:** 56-60.

Nadel L, Moscovitch M (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology* **7:** 217-227.

Nader K (2003). Memory traces unbound. *Trends in Neurosciences* **26:** 65-72.

Nader K, Schafe GE, Le Doux JE (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* **406:** 722-726.

Nelson A, Killcross S (2006). Amphetamine exposure enhances habit formation. *Journal of Neuroscience* **26:** 3805-3812.

Newhouse PA, Potter A, Singh A (2004). Effects of nicotinic stimulation on cognitive performance. *Current Opinion in Pharmacology* **4:** 36-46.

NHSDA (2001). Summary of Findings from the 2000 National Household Survey on Drug Abuse [DHHS Publication Number (SMA) 01-3549]. Substance Abuse and Mental Health Services Administration (Department of Health and Human Services, USA).

Niv Y, Duff MO, Dayan P (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions* **1:** 6.

O'Brien CP (1997). A range of research-based pharmacotherapies for addiction. *Science* **278:** 66-70.

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

O'Brien CP, McLellan AT (1996). Myths about the treatment of addiction. *Lancet* **347:** 237-240.

O'Brien CP, O'Brien TJ, Mintz J, Brady JP (1975). Conditioning of narcotic abstinence symptoms in human subjects. *Drug and Alcohol Dependence* **1:** 115-123.

O'Brien CP, Testa T, O'Brien TJ, Brady JP, Wells B (1977). Conditioned narcotic withdrawal in humans. *Science* **195:** 1000-1002.

O'Brien CP, Testa T, O'Brien TJ, Greenstein R (1976). Conditioning in human opiate addicts. *Pavlov J Biol Sci* **11:** 195-202.

O'Carroll RE, Papps BP (2003). Decision making in humans: the effect of manipulating the central noradrenergic system. *Journal of Neurology, Neurosurgery and Psychiatry* **74:** 376-378.

Odling-Smee FJ (1975). Background stimuli and the inter-stimulus interval during Pavlovian conditioning. *Quarterly Journal of Experimental Psychology* **27:** 387-392.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304:** 452-454.

O'Doherty JP (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current Opinion in Neurobiology* **14:** 769-776.

O'Donnell P, Grace AA (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *Journal of Neuroscience* **15:** 3622-3639.

Ohmura Y, Takahashi T, Kitamura N (2005). Discounting delayed and probabilistic monetary gains and losses by smokers of cigarettes. *Psychopharmacology* **182:** 508-515.

O'Keefe J, Dostrovsky J (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* **34:** 171-175.

O'Keefe J, Nadel L (1978). *The Hippocampus as a Cognitive Map*. Oxford, New York.

Olausson P, Jentsch JD, Taylor JR (2003). Repeated Nicotine Exposure Enhances Reward-Related Learning in the Rat. *Neuropsychopharmacology* **28:** 1264-1271.

Olds J, Milner P (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology* **47:** 419-427.

Olmstead MC, Lafond MV, Everitt BJ, Dickinson A (2001). Cocaine seeking by rats is a goal-directed action. *Behavioral Neuroscience* **115:** 394-402.

Olton DS, Meck WH, Church RM (1987). Separation of hippocampal and amygdaloid involvement in temporal memory dysfunctions. *Brain Research* **404:** 180-188.

Ortner CN, MacDonald TK, Olmstead MC (2003). Alcohol intoxication reduces impulsivity in the delay-discounting paradigm. *Alcohol and Alcoholism* **38:** 151-156.

Ostaszewski P, Green L, Myerson J (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychonomic Bulletin & Review* **5:** 324-333.

Ostaszewski P, Karzel K (2002). Discounting of delayed and probabilistic losses of different amounts. *European Psychologist* **7:** 295-301.

OUP (1997). *New Shorter Oxford English Dictionary*. Oxford University Press, Oxford, UK.

Packard MG, 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.

Padoa-Schioppa C, Assad JA (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* **441:** 223-226.

Paine TA, Dringenberg HC, Olmstead MC (2003). Effects of chronic cocaine on impulsivity: relation to cortical serotonin mechanisms. *Behavioural Brain Research* **147:** 135-147.

Papa M, Sagvolden T, Sergeant JA, 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, 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, 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.

Papez JW (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry* **38:** 725-743.

Parkinson JA (1998). Limbic corticostriatal circuitry underlying Pavlovian associative learning. Unpublished PhD thesis, University of Cambridge.

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

Parkinson JA, Dalley JW, Cardinal RN, Bamford A, Fehnert B, Lachenal G, Rudarakanchana N, Halkerston KM, Robbins TW, Everitt BJ (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behavioural Brain Research* **137:** 149-163.

Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ (1999a). Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive Pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by d-amphetamine. *Journal of Neuroscience* **19:** 2401-2411.

Parkinson JA, Robbins TW, Everitt BJ (1999b). 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, Robbins TW, Everitt BJ (2000b). Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. *European Journal of Neuroscience* **12:** 405-413.

Parkinson JA, Willoughby PJ, Robbins TW, Everitt BJ (2000c). 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.

Parrott S, Godfrey C, Raw M, West R, McNeill A (1998). Guidance for commissioners on the cost effectiveness of smoking cessation interventions. Health Educational Authority. *Thorax* **53 Suppl 5 Pt 2:** S1-38.

Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* **19:** 1439-1448.

Paus T (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature Reviews Neuroscience* **2:** 417-424.

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

Paxinos G, Watson C (1997). *The Rat Brain in Stereotaxic Coordinates*. Compact Third (CD-Rom) edition. Academic Press.

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

Pearce JM, Roberts AD, Good M (1998). Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors. *Nature* **396:** 75-77.

Pecina S, Berridge KC, Parker LA (1997). Pimozide does not shift palatability: Separation of anhedonia from sensorimotor suppression by taste reactivity. *Pharmacology Biochemistry and Behavior* **58:** 801-811.

Pecina S, Cagniard B, Berridge KC, Aldridge JW, Zhuang X (2003). Hyperdopaminergic mutant mice have higher "wanting" but not "liking" for sweet rewards. *Journal of Neuroscience* **23:** 9395-9402.

Pecina S, Schulkin J, Berridge KC (2006). Nucleus accumbens corticotropin-releasing factor increases cue-triggered motivation for sucrose reward: paradoxical positive incentive effects in stress? *BMC Biol* **4:** 8.

Pennartz CM, Groenewegen HJ, Lopes da Silva FH (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Progress in Neurobiology* **42:** 719-761.

Perin CT (1943). A quantitative investigation of the delay-of-reinforcement gradient. *Journal of Experimental Psychology* **32:** 37-51.

Petry NM (2001). Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology* **154:** 243-250.

Petry NM, Bickel WK, Arnett M (1998). Shortened time horizons and insensitivity to future consequences in heroin addicts. *Addiction* **93:** 729-738.

Pettit HO, Ettenberg A, Bloom FE, Koob GF (1984). Destruction of dopamine in the nucleus accumbens selectively attenuates cocaine but not heroin self-administration in rats. *Psychopharmacology* **84:** 167-173.

Pettit HO, Justice JB, Jr. (1989). Dopamine in the nucleus accumbens during cocaine self-administration as studied by in vivo microdialysis. *Pharmacology, Biochemistry and Behavior* **34:** 899-904.

Pettit HO, Pan HT, Parsons LH, Justice JB, Jr. (1990). Extracellular concentrations of cocaine and dopamine are enhanced during chronic cocaine administration. *Journal of Neurochemistry* **55:** 798-804.

Phillips AG, Ahn S, Howland JG (2003). Amygdalar control of the mesocorticolimbic dopamine system: parallel pathways to motivated behavior. *Neuroscience and Biobehavioral Reviews* **27:** 543-554.

Phillips GD, Robbins TW, Everitt BJ (1994). Bilateral intra-accumbens self-administration of d-amphetamine: antagonism with intra-accumbens SCH-23390 and sulpiride. *Psychopharmacology* **114:** 477-485.

Phillips RG, LeDoux JE (1992). Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behavioral Neuroscience* **106:** 274-285.

Phillips RG, LeDoux JE (1994). Lesions of the dorsal hippocampal formation interfere with background but not foreground contextual fear conditioning. *Learning & Memory* **1:** 34-44.

Phillips RG, LeDoux JE (1995). Lesions of the fornix but not the entorhinal or perirhinal cortex interfere with contextual fear conditioning. *Journal of Neuroscience* **15:** 5308-5315.

Philpot RM, McQuown S, Kirstein CL (2001). Stereotaxic localization of the developing nucleus accumbens septi. *Brain Research. Developmental Brain Research* **130:** 149-153.

Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience* **23:** 11078-11084.

Pilla M, Perachon S, Sautel F, Garrido F, Mann A, Wermuth CG, Schwartz JC, Everitt BJ, Sokoloff P (1999). Selective inhibition of cocaine-seeking behaviour by a partial dopamine D3 receptor agonist. *Nature* **400:** 371-375.

Pitkänen A (2000). Connectivity of the rat amygdaloid complex. In *The amygdala: a functional analysis*, Second edition (Aggleton JP, ed.), pp. 31-115. Oxford University Press, New York.

Plutchik R, Van Praag H (1989). The measurement of suicidality, aggressivity and impulsivity. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **13 Suppl:** S23-34.

Port R, Curtis K, Inoue C, Briggs J, Seybold K (1993). Hippocampal damage does not impair instrumental appetitive conditioning with delayed reinforcement. *Brain Research Bulletin* **30:** 41-44.

Port RL, Romano AG, Steinmetz JE, Mikhail AA, Patterson MM (1986). Retention and acquisition of classical trace conditioned responses by rabbits with hippocampal lesions. *Behavioral Neuroscience* **100:** 745-752.

Post RM, Weiss SR (1988). Psychomotor stimulant vs. local anesthetic effects of cocaine: role of behavioral sensitization and kindling. *NIDA Research Monograph* **88:** 217-238.

Pothuizen HH, Jongen-Relo AL, Feldon J, Yee BK (2005). Double dissociation of the effects of selective nucleus accumbens core and shell lesions on impulsive-choice behaviour and salience learning in rats. *European Journal of Neuroscience* **22:** 2605-2616.

Pothuizen HH, Jongen-Relo AL, Feldon J, Yee BK (2006). Latent inhibition of conditioned taste aversion is not disrupted, but can be enhanced, by selective nucleus accumbens shell lesions in rats. *Neuroscience* **137:** 1119-1130.

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

Poulos CX, Parker JL, 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.

Pu L, Bao GB, Xu NJ, Ma L, Pei G (2002). Hippocampal long-term potentiation is reduced by chronic opiate treatment and can be restored by re-exposure to opiates. *Journal of Neuroscience* **22:** 1914-1921.

Quirk GJ, Likhtik E, Pelletier JG, Pare D (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience* **23:** 8800-8807.

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

Rachlin H (1997). Four teleological theories of addiction. *Psychonomic Bulletin & Review* **4:** 462-473.

Rachlin H (2000a). The lonely addict. In *Reframing health behavior change with behavioral economics* (Bickel WK, Vuchinich RE, eds.), pp. 145-166. Lawrence Erlbaum Associates, Mahwah, NJ.

Rachlin H (2000b). *The science of self-control*. Harvard University Press, Cambridge, Massachusetts.

Rachlin H (2003). Economic concepts in the behavioral study of addiction. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 129-149. Elsevier, Oxford.

Rachlin H, Brown J, Cross D (2000). Discounting in judgments of delay and probability. *Journal of Behavioral Decision Making* **13:** 145-149.

Rachlin H, Castrogiovanni A, Cross D (1987). Probability and delay in commitment. *Journal of the Experimental Analysis of Behavior* **48:** 347-353.

Rachlin H, Green L (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior* **17:** 15-22.

Rachlin H, Green L, Kagel J, Battalio R (1976). Economic demand theory and psychological studies of choice. In *The Psychology of Learning and Motivation* (Bower G, ed.), pp. 129-154. Academic Press, New York.

Rachlin H, Logue AW, Gibbon J, Frankel M (1986). Cognition and behavior in studies of choice. *Psychological Review* **93:** 33-45.

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

Rachlin H, Siegel E (1994). Temporal patterning in probabilistic choice. *Organizational Behavior and Human Decision Processes* **59:** 161-176.

Rahman S, Sahakian BJ, Cardinal RN, Rogers RD, Robbins TW (2001). Decision making and neuropsychiatry. *Trends in Cognitive Sciences* **5:** 271-277.

Rawlins JN, Feldon J, Butt S (1985). The effects of delaying reward on choice preference in rats with hippocampal or selective septal lesions. *Behavioural Brain Research* **15:** 191-203.

Rawlins JN, Winocur G, Gray JA (1983). The hippocampus, collateral behavior, and timing. *Behavioral Neuroscience* **97:** 857-872.

Reading PJ, Dunnett SB (1991). The effects of excitotoxic lesions of the nucleus accumbens on a matching to position task. *Behavioural Brain Research* **46:** 17-29.

Reading PJ, Dunnett SB (1995). Embryonic striatal grafts reverse the disinhibitory effects of ibotenic acid lesions of the ventral striatum. *Experimental Brain Research* **105:** 76-86.

Reilly S, Good M (1989). Hippocampal lesions and associative learning in the pigeon. *Behavioral Neuroscience* **103:** 731-742.

Rescorla RA (1982). Effect of a stimulus intervening between CS and US in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes* **8:** 131-141.

Rescorla RA (1990a). Evidence for an association between the discriminative stimulus and the response-outcome association in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes* **16:** 326-334.

Rescorla RA (1990b). The role of information about the response-outcome relation in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes* **16:** 262-270.

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

Reynolds B (2006). The Experiential Discounting Task is sensitive to cigarette-smoking status and correlates with a measure of delay discounting. *Behavioural Pharmacology* **17:** 133-142.

Reynolds B, Richards JB, Dassinger M, de Wit H (2004a). Therapeutic doses of diazepam do not alter impulsive behavior in humans. *Pharmacology, Biochemistry and Behavior* **79:** 17-24.

Reynolds B, Richards JB, de Wit H (2006). Acute-alcohol effects on the Experiential Discounting Task (EDT) and a question-based measure of delay discounting. *Pharmacology, Biochemistry and Behavior*.

Reynolds B, Richards JB, Horn K, Karraker K (2004b). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behavioural Processes* **65:** 35-42.

Reynolds JN, Hyland BI, Wickens JR (2001). A cellular mechanism of reward-related learning. *Nature* **413:** 67-70.

Reynolds JN, Wickens JR (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks* **15:** 507-521.

Richards JB, Chock MA, Carlson B, de Wit H, 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, 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, de Wit H (1999a). Effects of methamphetamine on the adjusting amount procedure, a model of impulsive behavior in rats. *Psychopharmacology* **146:** 432-439.

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

Richards JB, Zhang L, Mitchell SH, de Wit H (1999b). Delay or probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the Experimental Analysis of Behavior* **71:** 121-143.

Richardson NR, 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.

Richmond J, Colombo M (2002). Hippocampal lesions, contextual retrieval, and autoshaping in pigeons. *Brain Research* **928:** 60-68.

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, Cardinal RN, DiCiano P, Halligan PW, Hellemans KGC, Lee JLC, Everitt BJ (2005). Neuroscience of drugs and addiction [Foresight: Brain Science, Addiction and Drugs project; www.foresight.gov.uk]. UK Office of Science and Technology.

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

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

Robbins TW, Everitt BJ (1999). Drug addiction: bad habits add up [news]. *Nature* **398:** 567-570.

Robbins TW, Watson BA, Gaskin M, 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, Berridge KC (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews* **18:** 247-291.

Rogers RD, Everitt BJ, Baldacchino A, Blackshaw AJ, Swainson R, Wynne K, Baker NB, Hunter J, Carthy T, Booker E, London M, Deakin JF, Sahakian BJ, Robbins TW (1999a). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology* **20:** 322-339.

Rogers RD, Lancaster M, Wakeley J, Bhagwagar Z (2004a). Effects of beta-adrenoceptor blockade on components of human decision-making. *Psychopharmacology* **172:** 157-164.

Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, Robbins TW (1999b). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience* **19:** 9029-9038.

Rogers RD, Ramnani N, Mackay C, Wilson JL, Jezzard P, Carter CS, Smith SM (2004b). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry* **55:** 594-602.

Rogers RD, Tunbridge EM, Bhagwagar Z, Drevets WC, Sahakian BJ, Carter CS (2003). Tryptophan depletion alters the decision-making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology* **28:** 153-162.

Rolls ET (2000). Hippocampo-cortical and cortico-cortical backprojections. *Hippocampus* **10:** 380-388.

Rosenbaum RS, Winocur G, Moscovitch M (2001). New views on old memories: re-evaluating the role of the hippocampal complex. *Behavioural Brain Research* **127:** 183-197.

Rosenkranz JA, Moore H, Grace AA (2003). The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *Journal of Neuroscience* **23:** 11054-11064.

Rosenzweig ES, Barnes CA, McNaughton BL (2002). Making room for new memories. *Nature Neuroscience* **5:** 6-8.

Rotter JB (1954). *Social learning and clinical psychology*. Prentice-Hall, Englewood Cliffs, NJ.

Roy A, DeJong J, Linnoila M (1989). Extraversion in pathological gamblers: correlates with indices of noradrenergic function. *Archives of General Psychiatry* **46:** 679-681.

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

Rubin RD (1976). Clinical use of retrograde amnesia produced by electroconvulsive shock. A conditioning hypothesis. *Can Psychiatr Assoc J* **21:** 87-90.

Rubin RD, Fried R, Franks CM (1969). New application of ECT. In *Advances in Behavior Therapy, 1968* (Rubin RD, Franks CM, eds.), pp. 37-44. Academic Press, New York.

Rudy JW, Barrientos RM, O'Reilly RC (2002). Hippocampal formation supports conditioning to memory of a context. *Behavioral Neuroscience* **116:** 530-538.

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

Russell V, de Villiers A, Sagvolden T, Lamm M, 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, 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.

Russell VA, Sagvolden T, Johansen EB (2005). Animal models of attention-deficit hyperactivity disorder. *Behavioral and Brain Functions* **1:** 9.

Sacchetti B, Lorenzini CA, Baldi E, Tassoni G, Bucherelli C (1999). Auditory thalamus, dorsal hippocampus, basolateral amygdala, and perirhinal cortex role in the consolidation of conditioned freezing to context and to acoustic conditioned stimulus in the rat. *Journal of Neuroscience* **19:** 9570-9578.

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.

Saffer H, Chaloupka FJ (1995). The demand for illicit drugs [Working Paper 5238]. National Bureau of Economic Research.

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, 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, 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, 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, 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, Correa M (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research* **137:** 3-25.

Salamone JD, Correa M, Mingote SM, Weber SM (2003). Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. *Journal of Pharmacology and Experimental Therapeutics* **305:** 1-8.

Salamone JD, Correa M, Mingote SM, Weber SM (2005). Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Current Opinion in Pharmacology* **5:** 34-41.

Salamone JD, Cousins MS, Bucher S (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural Brain Research* **65:** 221-229.

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

Salamone JD, Wisniecki A, Carlson BB, Correa M (2001). Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience* **105:** 863-870.

Salinas JA, Introini-Collison IB, Dalmaz C, McGaugh JL (1997). Posttraining intraamygdala infusions of oxotremorine and propranolol modulate storage of memory for reductions in reward magnitude. *Neurobiology of Learning and Memory* **68:** 51-59.

Salinas JA, McGaugh JL (1996). The amygdala modulates memory for changes in reward magnitude: involvement of the amygdaloid GABAergic system. *Behavioural Brain Research* **80:** 87-98.

Salinas JA, Packard MG, McGaugh JL (1993). Amygdala modulates memory for changes in reward magnitude: reversible post-training inactivation with lidocaine attenuates the response to a reduction in reward. *Behavioural Brain Research* **59:** 153-159.

Savage LM, Buzzetti RA, Ramirez DR (2004). The effects of hippocampal lesions on learning, memory, and reward expectancies. *Neurobiology of Learning and Memory* **82:** 109-119.

Sax KW, Strakowski SM (2001). Behavioral sensitization in humans. *Journal of Addictive Diseases* **20:** 55-65.

Schaler JA (2000). *Addiction is a choice*. Open Court Publishing, Chicago, Illinois.

Schmaltz LW, Isaacson RL (1967). Effect of bilateral hippocampal destruction on the acquisition and extinction of an operant response. *Physiology and Behavior* **2:** 291-298.

Schmitt WB, Arianpour R, Deacon RM, Seeburg PH, Sprengel R, Rawlins JN, Bannerman DM (2004). The role of hippocampal glutamate receptor-A-dependent synaptic plasticity in conditional learning: the importance of spatiotemporal discontiguity. *Journal of Neuroscience* **24:** 7277-7282.

Schneider F, Habel U, Wagner M, Franke P, Salloum JB, Shah NJ, Toni I, Sulzbach C, Honig K, Maier W, Gaebel W, Zilles K (2001). Subcortical correlates of craving in recently abstinent alcoholic patients. *American Journal of Psychiatry* **158:** 1075-1083.

Schneiderman N (1966). Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning. *Journal of Comparative and Physiological Psychology* **62:** 397-402.

Schramm NL, Egli RE, Winder DG (2002). LTP in the mouse nucleus accumbens is developmentally regulated. *Synapse* **45:** 213-219.

Schultz W (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* **80:** 1-27.

Schultz W (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology* **14:** 139-147.

Schultz W (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology* **57:** 87-115.

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

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

Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* **275:** 1593-1599.

Schultz W, Dickinson A (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience* **23:** 473-500.

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

Schultz W, Tremblay L, Hollerman JR (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex* **10:** 272-284.

Scoville WB, Milner B (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry* **20:** 11-21.

Seeman P, Madras B (2002). Methylphenidate elevates resting dopamine which lowers the impulse-triggered release of dopamine: a hypothesis. *Behavioural Brain Research* **130:** 79-83.

Selden NR, Everitt BJ, Jarrard LE, Robbins TW (1991). Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience* **42:** 335-350.

Sell LA, Morris JS, Bearn J, Frackowiak RS, Friston KJ, Dolan RJ (2000). Neural responses associated with cue evoked emotional states and heroin in opiate addicts. *Drug and Alcohol Dependence* **60:** 207-216.

Shahan TA, Lattal KA (1998). On the functions of the changeover delay. *Journal of the Experimental Analysis of Behavior* **69:** 141-160.

Shastri L, Ajjanagadde V (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences* **16:** 417-494.

Shaw B (2000). UK Drug Policy Report to the Cleveland Police Authority. Cleveland Police.

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

Shors TJ (2004). Memory traces of trace memories: neurogenesis, synaptogenesis and awareness. *Trends in Neurosciences* **27:** 250-256.

Shors TJ, Townsend DA, Zhao M, Kozorovitskiy Y, Gould E (2002). Neurogenesis may relate to some but not all types of hippocampal-dependent learning. *Hippocampus* **12:** 578-584.

Shull RL, Pliskoff SS (1967). Changeover delay and concurrent schedules: some effects on relative performance measures. *Journal of the Experimental Analysis of Behavior* **10:** 517-527.

Siegel S (1975). Evidence from rats that morphine tolerance is a learned response. *Journal of Comparative and Physiological Psychology* **89:** 498-506.

Siegel S (1976). Morphine analgesic tolerance: its situation specificity supports a Pavlovian conditioning model. *Science* **193:** 323-325.

Siegel S (1988). Drug anticipation and the treatment of dependence. *NIDA Research Monograph* **84:** 1-24.

Siegel S (1999). Drug anticipation and drug addiction. The 1998 H. David Archibald Lecture. *Addiction* **94:** 1113-1124.

Skinner BF (1938). *The behavior of organisms: an experimental analysis*. Appleton, New York.

Skinner BF (1953). *Science and Human Behavior*. Macmillan, New York.

Skog O-J (2003). Addiction: definitions and mechanisms. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 157-175, 182. Elsevier, Oxford.

Slovic P, Fischhoff B, Lichtenstein S (1982). Facts versus fears: understanding perceived risk. In *Judgement Under Uncertainty: Heuristics and Biases* (Kahneman D, Slovic P, Tversky A, eds.), pp. 463-489. Cambridge University Press, New York.

Slovic P, Lichtenstein S (1983). Preference reversals: a broader perspective. *American Economic Review* **73:** 596-605.

Smith DM, Mizumori SJ (2006). Learning-related development of context-specific neuronal responses to places and events: the hippocampal role in context processing. *Journal of Neuroscience* **26:** 3154-3163.

Smith Z (1999). The revenue effect of changing alcohol duties. Institute for Fiscal Studies.

Smith-Roe SL, Kelley AE (2000). Coincident activation of NMDA and dopamine D1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. *Journal of Neuroscience* **20:** 7737-7742.

Sokoloff P, Giros B, Martres MP, Bouthenet ML, Schwartz JC (1990). Molecular cloning and characterization of a novel dopamine receptor (D3) as a target for neuroleptics. *Nature* **347:** 146-151.

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

Solanto MV (2002). Dopamine dysfunction in AD/HD: integrating clinical and basic neuroscience research. *Behavioural Brain Research* **130:** 65-71.

Solomon PR, Vander Schaaf ER, Thompson RF, Weisz DJ (1986). Hippocampus and trace conditioning of the rabbit's classically conditioned nictitating membrane response. *Behavioral Neuroscience* **100:** 729-744.

Solomon RL (1980a). The opponent-process theory of acquired motivation: the costs of pleasure and the benefits of pain. *American Psychologist* **35:** 691-712.

Solomon RL (1980b). Recent experiments testing an opponent-process theory of acquired motivation. *Acta Neurobiol Exp (Wars)* **40:** 271-289.

Solomon RL, Corbit JD (1973). An opponent-process theory of motivation. II. Cigarette addiction. *Journal of Abnormal Psychology* **81:** 158-171.

Solomon RL, Corbit JD (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychological Review* **81:** 119-145.

Sonuga-Barke EJ (2002). Psychological heterogeneity in AD/HD--a dual pathway model of behaviour and cognition. *Behavioural Brain Research* **130:** 29-36.

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

Sougné J (1998). Connectionism and the problem of multiple instantiation. *Trends in Cognitive Sciences* **2:** 183-189.

Sozou PD (1998). On hyperbolic discounting and uncertain hazard rates [DOI: 10.1098/rspb.1998.0534]. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265:** 2015-2020.

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

Springer AD, Miller RR (1972). Retrieval failure induced by electroconvulsive shock: reversal with dissimilar training and recovery agents. *Science* **177:** 628-630.

Squire LR (1980). Specifying the defect in human amnesia: storage, retrieval and semantics. *Neuropsychologia* **18:** 369-372.

Squire LR (1986). Mechanisms of memory. *Science* **232:** 1612-1619.

Squire LR (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review* **99:** 195-231.

Squire LR, Clark RE, Knowlton BJ (2001). Retrograde amnesia. *Hippocampus* **11:** 50-55.

Squire LR, Davis HP, Spanis CW (1980). Neurobiology of amnesia. *Science* **209:** 836-837.

Squire LR, Shimamura AP, Graf P (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia* **25:** 195-210.

Squire LR, Slater PC, Chace PM (1975). Retrograde amnesia: temporal gradient in very long term memory following electroconvulsive therapy. *Science* **187:** 77-79.

Stevens KE, Shiotsu G, Stein L (1991). Hippocampal mu-receptors mediate opioid reinforcement in the CA3 region. *Brain Research* **545:** 8-16.

Stevenson MK (1986). A discounting model for decisions with delayed positive and negative outcomes. *Journal of Experimental Psychology: General* **115:** 131-154.

Stigler G, Becker GS (1977). De gustibus non est disputandum. *American Economic Review* **67:** 76-90.

Stockmeier CA (2003). Involvement of serotonin in depression: evidence from postmortem and imaging studies of serotonin receptors and the serotonin transporter. *Journal of Psychiatric Research* **37:** 357-373.

Stratford TR, Kelley AE (1997). GABA in the nucleus accumbens shell participates in the central regulation of feeding behavior. *Journal of Neuroscience* **17:** 4434-4440.

Streatfeild D (2001). *Cocaine*. Virgin, London.

Stricker EM, Zigmond MJ (1976). Recovery of function after damage to central catecholamine-containing neurons: a neurochemical model for the lateral hypothalamic syndrome. In *Progress in Psychology and Physiological Psychology* (Sprague JM, Epstein AN, eds.), pp. 121-188. Academic Press, New York.

Stroustrup B (1986). *The C++ Programming Language*. Addison-Wesley, Reading, Massachusetts.

Sullivan JM, Zhang JQ (2005). MIVA 0.9 [http://ccni.wpi.edu]. Centre for Comparative Neuroimaging.

Sutton RS (1988). Learning to predict by the method of temporal differences. *Machine Learning* **3:** 9-44.

Swanson CJ, Heath S, Stratford TR, Kelley AE (1997). Differential behavioral responses to dopaminergic stimulation of nucleus accumbens subregions in the rat. *Pharmacology Biochemistry and Behavior* **58:** 933-945.

Swanson J, Castellanos FX, Murias M, LaHoste G, Kennedy J (1998). Cognitive neuroscience of attention deficit hyperactivity disorder and hyperkinetic disorder. *Current Opinion in Neurobiology* **8:** 263-271.

Szerman N, Peris L, Mesias B, Colis P, Rosa J, Prieto A (2005). Reboxetine for the treatment of patients with Cocaine Dependence Disorder. *Hum Psychopharmacol* **20:** 189-192.

Takehara K, Kawahara S, Takatsuki K, Kirino Y (2002). Time-limited role of the hippocampus in the memory for trace eyeblink conditioning in mice. *Brain Research* **951:** 183-190.

Takikawa Y, Kawagoe R, Hikosaka O (2002). Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. *Journal of Neurophysiology* **87:** 508-515.

Taylor JR, Horger BA (1999). Enhanced responding for conditioned reward produced by intra-accumbens amphetamine is potentiated after cocaine sensitization. *Psychopharmacology* **142:** 31-40.

Taylor JR, Robbins TW (1984). Enhanced behavioural control by conditioned reinforcers following microinjections of d-amphetamine into the nucleus accumbens. *Psychopharmacology* **84:** 405-412.

Taylor JR, Robbins TW (1986). 6-Hydroxydopamine lesions of the nucleus accumbens, but not of the caudate nucleus, attenuate enhanced responding with reward-related stimuli produced by intra-accumbens d-amphetamine. *Psychopharmacology* **90:** 390-397.

Thiebot MH, Lebihan C, Soubrie P, Simon P (1985). Benzodiazepines Reduce the Tolerance to Reward Delay in Rats. *Psychopharmacology* **86:** 147-152.

Thompson AM, Swant J, Gosnell BA, Wagner JJ (2004). Modulation of long-term potentiation in the rat hippocampus following cocaine self-administration. *Neuroscience* **127:** 177-185.

Thorndike EL (1905). *The Elements of Psychology*. Seiler, New York.

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

Tiffany ST, Carter BL (1998). Is craving the source of compulsive drug use? *Journal of Psychopharmacology* **12:** 23-30.

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

Toates F (1986). *Motivational systems*. Cambridge University Press, Cambridge.

Tobler PN, Fiorillo CD, Schultz W (2005). Adaptive coding of reward value by dopamine neurons. *Science* **307:** 1642-1645.

Tolman EC (1932). *Purposive Behavior in Animals and Men*. Century, New York.

Tomie A, Aguado AS, Pohorecky LA, Benjamin D (1998). Ethanol induces impulsive-like responding in a delay-of-reward operant choice procedure: impulsivity predicts autoshaping. *Psychopharmacology* **139:** 376-382.

Träskman-Bendz L, Åsberg M, Schalling D (1986). Serotonergic function and suicidal behavior in personality disorders. *Annals of the New York Academy of Sciences* **487:** 168-174.

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

Uchtenhagen A (1997). Summary of the Synthesis Report. Institute for Social and Preventive Medicine at the University of Zurich.

Uhlenhuth EH, Johanson CE, Kilgore K, Kobasa SC (1981). Drug preference and mood in humans: preference for d-amphetamine and subject characteristics. *Psychopharmacology* **74:** 191-194.

UK (2000). *Spending Review 2000: New Public Spending Plans 2001-2004*. HM Treasury, London.

UK (2003). Alcohol misuse: how much does it cost? Cabinet Office Strategy Unit.

Ungerstedt U (1971). Stereotaxic mapping of the monoamine pathways in the rat brain. *Acta Physiologica Scandinavica* **82 (Suppl. 367):** 1-48.

Ursin R, Ursin H, Olds J (1966). Self-stimulation of hippocampus in rats. *Journal of Comparative and Physiological Psychology* **61:** 353-359.

USA (2001). Leading Causes Charts [National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; www.cdc.gov/ncipc/osp/charts.htm].

Uylings HB, Groenewegen HJ, Kolb B (2003). Do rats have a prefrontal cortex? *Behavioural Brain Research* **146:** 3-17.

Uz T, Ahmed R, Akhisaroglu M, Kurtuncu M, Imbesi M, Dirim Arslan A, Manev H (2005). Effect of fluoxetine and cocaine on the expression of clock genes in the mouse hippocampus and striatum. *Neuroscience* **134:** 1309-1316.

Vanderschuren LJ, Everitt BJ (2004). Drug seeking becomes compulsive after prolonged cocaine self-administration. *Science* **305:** 1017-1019.

Villarreal DM, Do V, Haddad E, Derrick BE (2002). NMDA receptor antagonists sustain LTP and spatial memory: active processes mediate LTP decay. *Nature Neuroscience* **5:** 48-52.

Vogt BA (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience* **6:** 533-544.

Volkow ND, Fowler JS, Wang GJ (1999). Imaging studies on the role of dopamine in cocaine reinforcement and addiction in humans. *Journal of Psychopharmacology* **13:** 337-345.

Volkow ND, Wang GJ, Fowler JS (1997). Imaging studies of cocaine in the human brain and studies of the cocaine addict. *Annals of the New York Academy of Sciences* **820:** 41-54; discussion 54-45.

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

Vorel SR, Ashby CR, Jr., Paul M, Liu X, Hayes R, Hagan JJ, Middlemiss DN, Stemp G, Gardner EL (2002). Dopamine D3 receptor antagonism inhibits cocaine-seeking and cocaine-enhanced brain reward in rats. *Journal of Neuroscience* **22:** 9595-9603.

Vorel SR, Liu X, Hayes RJ, Spector JA, Gardner EL (2001). Relapse to cocaine-seeking after hippocampal theta burst stimulation. *Science* **292:** 1175-1178.

Vuchinich RE, Calamas ML (1997). Does the repeated gambles procedure measure impulsivity in social drinkers? *Experimental and Clinical Psychopharmacology* **5:** 157-162.

Vuchinich RE, Heather N (2003). Introduction: overview of behavioural economic perspectives on substance use and addiction. In *Choice, Behavioral Economics and Addiction* (Heather N, Vuchinich RE, eds.), pp. 1-31. Elsevier, Oxford.

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

Wagner AR, Logan FA, Haberlandt K (1968). Stimulus selection in animal discrimination learning. *J Exp Psychol* **76:** 177-186.

Wakabayashi KT, Fields HL, Nicola SM (2004). Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. *Behavioural Brain Research* **154:** 19-30.

Walderhaug E, Lunde H, Nordvik JE, Landro NI, Refsum H, Magnusson A (2002). Lowering of serotonin by rapid tryptophan depletion increases impulsiveness in normal individuals. *Psychopharmacology* **164:** 385-391.

Wallenstein GV, Eichenbaum H, Hasselmo ME (1998). The hippocampus as an associator of discontiguous events. *Trends in Neurosciences* **21:** 317-323.

Walton ME, Bannerman DM, Alterescu K, Rushworth MF (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *Journal of Neuroscience* **23:** 6475-6479.

Walton ME, Bannerman DM, Rushworth MF (2002). The role of rat medial frontal cortex in effort-based decision making. *Journal of Neuroscience* **22:** 10996-11003.

Walton ME, Croxson PL, Rushworth MF, Bannerman DM (2005). The mesocortical dopamine projection to anterior cingulate cortex plays no role in guiding effort-related decisions. *Behavioral Neuroscience* **119:** 323-328.

Ward HG, Nicklous DM, Aloyo VJ, Simansky KJ (2006). Mu-opioid receptor cellular function in the nucleus accumbens is essential for hedonically driven eating. *European Journal of Neuroscience* **23:** 1605-1613.

Warner LA, Kessler RC, Hughes M, Anthony JC, Nelson CB (1995). Prevalence and correlates of drug use and dependence in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry* **52:** 219-229.

Warrington EK, Weiskrantz L (1970). Amnesic syndrome: consolidation or retrieval? *Nature* **228:** 628-630.

Waterhouse JW (1891). Ulysses and the Sirens.

Watkins CJCH (1989). Learning from delayed rewards. Unpublished PhD thesis, University of Cambridge.

Weatherly JN, McSweeney FK, Swindell S (2004). Within-session rates of responding when reinforcer magnitude is changed within the session. *Journal of General Psychology* **131:** 5-16.

Weiss C, Bouwmeester H, Power JM, Disterhoft JF (1999a). Hippocampal lesions prevent trace eyeblink conditioning in the freely moving rat. *Behavioural Brain Research* **99:** 123-132.

Weiss C, Knuttinen MG, Power JM, Patel RI, O'Connor MS, Disterhoft JF (1999b). Trace eyeblink conditioning in the freely moving rat: optimizing the conditioning parameters. *Behavioral Neuroscience* **113:** 1100-1105.

Weiss RD, Mirin SM (1986). Subtypes of cocaine abusers. *Psychiatric Clinics of North America* **9:** 491-501.

White AM, Swartzwelder HS (2004). Hippocampal function during adolescence: a unique target of ethanol effects. *Annals of the New York Academy of Sciences* **1021:** 206-220.

White KG, 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.

Wickens J, Kötter R (1995). Cellular models of reinforcement. In *Models of Information Processing in the Basal Ganglia* (Houk JC, Davis JL, Beiser DG, eds.), pp. 187-214. MIT Press, Cambridge, Massachusetts / London.

Wikler A (1965). Conditioning factors in opiate addiction and relapse. In *Narcotics* (Wilner DI, Kessenbaum GG, eds.), pp. 85-100. McGraw-Hill, New York.

Wikler A (1973). Dynamics of drug dependence: Implications of a conditioning theory for research and treatment. *Archives of General Psychiatry* **28:** 611-616.

Will MJ, Franzblau EB, Kelley AE (2003). Nucleus accumbens mu-opioids regulate intake of a high-fat diet via activation of a distributed brain network. *Journal of Neuroscience* **23:** 2882-2888.

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

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

Williams DR, Williams H (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent nonreinforcement. *Journal of the Experimental Analysis of Behavior* **12:** 511-520.

Williams J, Dayan P (2005). Dopamine, learning, and impulsivity: a biological account of attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology* **15:** 160-179; discussion 157-169.

Williams WA, Shoaf SE, Hommer D, Rawlings R, Linnoila M (1999). Effects of acute tryptophan depletion on plasma and cerebrospinal fluid tryptophan and 5-hydroxyindoleacetic acid in normal volunteers. *Journal of Neurochemistry* **72:** 1641-1647.

Wilson C, Nomikos GG, Collu M, Fibiger HC (1995). Dopaminergic correlates of motivated behavior: importance of drive. *Journal of Neuroscience* **15:** 5169-5178.

Winstanley CA, Baunez C, Theobald DE, Robbins TW (2005a). Lesions to the subthalamic nucleus decrease impulsive choice but impair autoshaping in rats: the importance of the basal ganglia in Pavlovian conditioning and impulse control. *European Journal of Neuroscience* **21:** 3107-3116.

Winstanley CA, Dalley JW, Theobald DE, Robbins TW (2003). Global 5-HT depletion attenuates the ability of amphetamine to decrease impulsive choice on a delay-discounting task in rats. *Psychopharmacology* **170:** 320-331.

Winstanley CA, Dalley JW, Theobald DE, Robbins TW (2004a). Fractionating impulsivity: contrasting effects of central 5-HT depletion on different measures of impulsive behavior. *Neuropsychopharmacology* **29:** 1331-1343.

Winstanley CA, Theobald DE, Cardinal RN, Robbins TW (2004b). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *Journal of Neuroscience* **24:** 4718-4722.

Winstanley CA, Theobald DE, Dalley JW, Glennon JC, Robbins TW (2004c). 5-HT2A and 5-HT2C receptor antagonists have opposing effects on a measure of impulsivity: interactions with global 5-HT depletion. *Psychopharmacology* **176:** 376-385.

Winstanley CA, Theobald DE, Dalley JW, Robbins TW (2005b). Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* **30:** 669-682.

Winston GC (1980). Addiction and backsliding: a theory of compulsive consumption. *Journal of Economic Behavior and Organization* **1:** 295-324.

Wise RA (1981). Brain dopamine and reward. In *Theory in Psychopharmacology Volume 1* (Cooper SJ, ed.), 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 *Appetite: Neural and Behavioural Bases* (Legg CR, Booth D, eds.), pp. 243-263. Oxford University Press, New York.

Wise RA (1996). Neurobiology of addiction. *Current Opinion in Neurobiology* **6:** 243-251.

Wise RA, Bozarth MA (1985). Brain mechanisms of drug reward and euphoria. *Psychiatric Medicine* **3:** 445-460.

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

Wood ER, Dudchenko PA, Eichenbaum H (1999). The global record of memory in hippocampal neuronal activity. *Nature* **397:** 613-616.

Wultz B, Sagvolden T, Moser EI, 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, 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, Berridge KC (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered "wanting" for sucrose reward. *Journal of Neuroscience* **21:** 7831-7840.

Yamaguchi M, Suzuki T, Seki T, Namba T, Juan R, Arai H, Hori T, Asada T (2004). Repetitive cocaine administration decreases neurogenesis in adult rat hippocampus. *Annals of the New York Academy of Sciences* **1025:** 351-362.

Yamaguchi M, Suzuki T, Seki T, Namba T, Liu J, Arai H, Hori T, Shiga T (2005). Decreased cell proliferation in the dentate gyrus of rats after repeated administration of cocaine. *Synapse* **58:** 63-71.

Yin HH, Knowlton BJ, Balleine BW (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience* **19:** 181-189.

Young AM (2004). Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. *Journal of Neuroscience Methods* **138:** 57-63.

Yu AJ, Dayan P (2005). Uncertainty, neuromodulation, and attention. *Neuron* **46:** 681-692.

Zaborszky L, Alheid GF, Beinfeld MC, Eiden LE, Heimer L, Palkovits M (1985). Cholecystokinin innervation of the ventral striatum - a morphological and radioimmunological study. *Neuroscience* **14:** 427.

Zahm DS, Brog JS (1992). On the significance of subterritories in the "accumbens" part of the rat ventral striatum. *Neuroscience* **50:** 751-767.

Zhang M, Gosnell BA, Kelley AE (1998). Intake of high-fat food is selectively enhanced by mu opioid receptor stimulation within the nucleus accumbens. *Journal of Pharmacology and Experimental Therapeutics* **285:** 908-914.

Zhang M, Kelley AE (2000). Enhanced intake of high-fat food following striatal mu-opioid stimulation: microinjection mapping and fos expression. *Neuroscience* **99:** 267-277.

Zhang M, Kelley AE (2002). Intake of saccharin, salt, and ethanol solutions is increased by infusion of a mu opioid agonist into the nucleus accumbens. *Psychopharmacology* **159:** 415-423.

Zhuang X, Oosting RS, Jones SR, Gainetdinov RR, Miller GW, Caron MG, Hen R (2001). Hyperactivity and impaired response habituation in hyperdopaminergic mice. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 1982-1987.

Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004). Human striatal responses to monetary reward depend on saliency. *Neuron* **42:** 509-517.

Zola-Morgan S, Squire LR, Amaral DG (1986). Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience* **6:** 2950-2967.

Index of first authors

Page references are given here to the first author of cited studies.

Q

R

S

Y

Z