Neurobiology of delayed reinforcement

Rudolf Cardinal
Department of Experimental Psychology, University of Cambridge

British Psychological Society Annual Conference, Imperial College London
10:10–10:50 am, Saturday 17 April 2004
Delayed reinforcement:
the problems.
How do animals succeed in bridging delays to reinforcement?

... actions are by no means always followed by their outcomes, especially at a neuronal timescale

• ‘Impulsivity’ refers to several, dissociable tendencies:
  • ‘preparation’ impulsivity — failure to collect sufficient information to make a good decision
  • ‘motor’/‘execution’ impulsivity — inability to restrain actions
  • ‘outcome’ impulsivity — impulsive choice — preference for immediate, small rewards over large, delayed rewards.

Why do some individuals exhibit abnormally impulsive choice, choosing small, immediate rewards over large, delayed rewards?

... can be considered a normal personality trait (Aristotle, 350 BC)

... but impulsive choice contributes to attention-deficit/hyperactivity disorder (ADHD), drug addiction, mania, and personality disorders
Learning with delayed reinforcement
Discrimination learning with delayed reinforcement

Grice (1948)

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 $\times 1000$ 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.
Free-operant learning with delayed reinforcement

Dickinson, Watt & Griffiths (1992)
Signalled and unsignalled delayed reinforcement

Unsignalled

Signalled

signal or cue bridging delay
Cues present during the delay speed up learning

Grice (1948)

Fig. 5. Learning curves for the three different groups with five sec. delay
Choice
with delayed reinforcement
Temporal discounting: devaluing the future
Would you rather have £20 now, or £40 next year? We can call it \textit{impulsive} to choose the smaller-sooner reward, and \textit{self-controlled} to choose the larger-later reward. Three guesses about why people are impulsive (Ainslie, 1975):

\begin{itemize}
  \item They lack insight into the consequences of their actions
  \item They are aware of the consequences of their actions, but are unable to suppress some lower principle ("the devil, repetition compulsion, classical conditioning")
  \item They are aware of the consequences of their actions, and choose rationally according to their value system, but their values are distorted so that imminent consequences have a greater weight than remote ones — reduced value of delayed reinforcement.
\end{itemize}

\textit{Ainslie (1975)}
Impulsive and self-controlled individuals discount differently.
Hyperbolic temporal discounting: irrational, but true

Exponential: \[ \text{value} = \text{magnitude} \times e^{-K \cdot \text{delay}} \]

Hyperbolic: \[ \text{value} = \frac{\text{magnitude}}{1 + K \cdot \text{delay}} \]
Choosing future rewards: preference reversal

Ainslie (1975)
Pre-commitment as a means of self-control

Homer (1700 BC?) *Odyssey*; Waterhouse (1891) *Ulysses and the Sirens*

Pigeons are often impulsive (Rachlin & Green, 1972) — but they too exhibit pre-commitment (Ainslie, 1974; Ainslie & Herrnstein, 1981).
Steeper temporal discounting in drug addicts

Smokers discount money more than non- or ex-smokers

Similarly for heroin addicts

Smokers discount cigarettes more than money

Bickel et al. (1999), smokers; Madden et al. (1999), heroin addicts
Neurochemistry of choice involving delayed reinforcement
Serotonin (5HT) in impulsive choice

- Low levels of 5HT metabolites in cerebrospinal fluid associated with impulsive aggression and violence in humans (Åsberg 1976; Linnoila et al. 1983).
- 5HT involved in inhibition of behaviour (impulse control)? (Soubrié 1986)
- Lower levels of 5HT metabolites in cerebrospinal fluid of macaques making longer/‘riskier’ leaps through forest canopy! (Mehlman et al. 1994)

In studies specifically of impulsive choice:

- However, not clear cut: global 5HT depletion or antagonists do not always promote impulsive choice (Evenden & Ryan 1996; Crean et al. 2002; Winstanley et al. 2003) and 5HT2 agonists promote impulsive choice (Evenden & Ryan, 1996).
Dopamine (DA) in impulsive choice

- Amphetamine and methylphenidate (Ritalin), catecholamine releasers and reuptake blockers, are effective therapies for ADHD (Bradley, 1937, and on).

- The spontaneously hypertensive rat, an animal model of ADHD, has abnormal DA systems (e.g. Russell et al. 1995). Hyperdopaminergic? Hypodopaminergic? Debated… (e.g. Zhuang et al. 2001, Seeman & Madras 2002).

- Is impulsivity in ADHD due to steeper ‘temporal discounting’, due to abnormal DA systems? (e.g. Sagvolden & Sergeant, 1998).

- D2 receptors promote choice of delayed rewards. The D2 antagonist raclopride and the D1+D2 antagonist flupenthixol decrease preference for delayed reinforcement; the D1 antagonist SCH23390 has no effect (Wade et al. 2000).

- The effects of psychostimulants are complex (pharmacologically and behaviourally)... do they promote
  - **self-controlled choice**? (Sagvolden ’92; Richards ’97/’99, Wade ’00, de Wit ’02)
  - **impulsive choice**? (Evenden & Ryan ’96; Charrier & Thiébot ’96; Logue ’92)
Choice involving delayed reinforcement: typical task

Signalled and unsignalled delayed reinforcement

"No cue" condition

"Cue" condition

stimulus light illuminated
The cue supports choice of the large, delayed reinforcer in rats trained in its presence

Cardinal et al. (2000)
Amphetamine cue-independently decreased preference for the delayed reward, but cue-dependently increased it.

Cardinal et al. (2000)

Resolves some contradictions.
Neuroanatomy of delayed reinforcement: (1) choice
Stereotaxic, excitotoxic lesions...

Carlson (1991)  
Cardinal et al. (2001)
Nucleus accumbens core (AcbC) lesions severely impaired the ability of rats to choose a delayed reward.

Cardinal et al. (2001)
… even in rats that exhibit very strong preference for the large reward when it is not delayed.

**Cardinal et al. (2003)**
Anterior cingulate cortex (ACC) lesions, which have been shown to produced ‘motor impulsivity’ in the 5-choice task, had no effect upon responding for delayed rewards

Cardinal et al. (2001)
... although they might affect ‘response effort’ choices

Walton et al. (2002, 2003)
Lesioned subjects chose the large reward less frequently at zero delay, and more frequently at long delays.

Medial prefrontal cortex (mPFC) lesions induced an insensitivity to the task contingencies.

Cardinal et al. (2001)

Lesions of the basolateral amygdala (BLA) make rats more *impulsive* in this task; lesions of the orbitofrontal cortex (OFC) make rats more *self-controlled*.

Redrawn from Winstanley et al. (in press 2004).

*Lesions made after training; no stimulus in delay; 1 (immediate) v. 4 (delayed) pellets.*
... but OFC lesions can also have the opposite effect!

Mobini et al. (2002). Lesions made before training; stimulus in delay; 1 versus 2 pellets.
Choice depends on reinforcer magnitude as well as delay...

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

- \( V \) value, \( d \) delay, \( q \) quantity
- \( K \) delay discounting parameter
- \( Q \) quantity discounting parameter

Orbitofrontal cortex (OFC) lesions affect both delay and magnitude discounting (Kheramin et al., 2002).

i.e. low values of \( Q \) (relative indifference between the two reinforcers, compared to normal) can also induce ‘impulsive’ choice.

After Bradshaw & Szabadi (1992); Ho et al. (1999); Kheramin et al. (2002)
Neuroanatomy of delayed reinforcement: (2) learning
Instrumental contingencies are harder to detect with a delay

Acquisition of free-operant instrumental responding on a fixed-ratio-1 schedule

a) Zero delay

lever presses

food pellets

b) 10- or 20-second delay

lever presses

food pellets

We’ve seen that nucleus accumbens core (AcbC) lesions impair choice of delayed reward. Is this because they can’t learn the contingency when reward is delayed?
Lesions of the AcbC again...

Cardinal & Cheung (unpublished)
AcbC lesions impair instrumental acquisition only when there is a delay between action and outcome (1)

*Cardinal & Cheung (unpublished)*
AcbC lesions impair instrumental acquisition only when there is a delay between action and outcome (2)

Cardinal & Cheung (unpublished)
Holds true even when experienced (rather than programmed) delays are examined

Cardinal & Cheung (unpublished)
What about magnitude discrimination? The matching ‘law’…

Two alternatives (e.g. levers) A and B. Both deliver reinforcement intermittently and somewhat unpredictable (e.g. variable interval schedule).

\[
\frac{R_A}{R_A + R_B} = \frac{r_A}{r_A + r_B}
\]

where \( R \) is response rate; \( r \) is (experienced) reinforcement rate

This should be a way of testing animals’ sensitivity to reinforcement magnitude… For example, if the two schedules deliver at the same rate but A delivers 1 pellet per reinforcement and B delivers 4 pellets per reinforcement, animals should allocate 80% of their responses to B.

Herrnstein (1961, 1970)
AcbC-lesioned rats *better* at magnitude discrimination?

Considerable undermatching (common: Williams, 1994). But shams and lesions were influenced by reinforcer allocation (lines not flat), and AcbC-lesioned rats were more influenced by this than shams (AcbC line has a significantly steeper gradient).

Consistent with studies using other techniques (e.g. Balleine & Killcross 1994, Brown & Bowman 1995).

So a ‘magnitude’ explanation can’t explain the effect of AcbC lesions to produce impulsive choice. Therefore, AcbC-lesioned rats must be specifically hypersensitive to the effects of delays.

*Cardinal & Cheung (unpublished)*
The ‘limbic’ corticostriatal circuit: delayed reinforcement

The diagram illustrates the connections between various brain regions involved in delayed reinforcement. Key components include:

- **Insula**
- **OFC**
- **mPFC**
- **ACC**
- **BLA**
- **CeA**
- **Acb core**
- **Acb shell**
- **MD thalamus**
- **VP**

Sensory information flows through these regions, with inputs from the brainstem and response systems. The diagram highlights the role of dopamine, serotonin, and acetylcholine in these processes. Different colors indicate the influence of these neurotransmitters:

- **Blue**: promotes self-controlled choice
- **Red**: promotes impulsive choice
- **Gray**: no influence?
- **No idea**: no specific influence indicated

The diagram also notes the involvement of various nuclei and projection systems, such as:

- VTA (dopamine)
- SNc (dopamine)
- Raphé nuclei (serotonin)
- Locus coeruleus (noradrenaline)
- Nucleus basalis (acetylcholine)

These components work together to modulate delayed reinforcement behaviors.
How to avoid temptation…

Pre-commitment strategies

Cues that signal the availability of the delayed outcome

Having a good amygdala/OFC/accumbens system to help you choose (and learn with) delayed rewards?

Waterhouse (1891)

Draper (1909)
Supervisor, collaborators and acknowledgements

- Barry Everitt
- Tim Cheung
- Caroline Parkinson
- John Parkinson
- David Pennicott
- Trevor Robbins
- Lakmali Sugathapala

- Helen Sweet-Gossage
- David Theobald
- Catharine Winstanley
- The UK Medical Research Council
- The Wellcome Trust
- The University of Cambridge School of Clinical Medicine