Chapter 5.
Local analysis of behaviour in the adjusting-delay task for assessing choice of delayed reinforcement

Abstract. The adjusting-delay task introduced by Mazur (1987) has been widely used to study choice of delayed reinforcers. The adjusting-delay paradigm involves repeated choice between one reinforcer delivered after a fixed delay and another, typically larger, reinforcer delivered after a variable delay; the variable delay is adjusted depending on the subject's choice until an equilibrium point is reached at which the subject is indifferent between the two alternatives. Rats were trained on a version of this task and their behaviour was examined to determine the nature of their sensitivity to the adjusting delay; these analyses included the use of a cross-correlational technique. No clear evidence of sensitivity to the adjusting delay was found. A number of decision rules, some sensitive to the adjusting delay and some not, were simulated to examine which effects usually supposed to be a consequence of delay sensitivity could be explained by delay-independent processes, such as a consistent, unchanging preference for one of the alternatives.

INTRODUCTION

While delayed reinforcement can have profound effects on learning (e.g. Grice, 1948; Dickinson *et al.*, 1992), it can also affect choice behaviour in well-trained animals. The effects of delays to reinforcement on choice have been extensively investigated in the consideration of 'impulsive choice' (Ainslie, 1975), exemplified by the inability of an individual to choose a large delayed reward in preference to a small immediate reward. As discussed in Chapter 1 (p. 60), choice with delayed reinforcement may be assessed using free-operant tasks such as the concurrent-chains procedure (Davison, 1987) or in discrete trials. Discrete-trial tasks may be further subdivided into 'systematic' tasks (e.g. Evenden & Ryan, 1996), in which the experimenter varies the delay to one or more of several reinforcers and then measures choice, and 'adjusting' tasks, in which the subject's behaviour determines which delays are to be sampled.

The adjusting-delay task was introduced by Mazur (1984; 1987; 1988). Its principle is as follows. Subjects are given repeated choices of a small reinforcer A delivered after a small fixed delay (dA, which may be zero to give immediate delivery) and a large reinforcer B delivered after a longer delay (dB). The delay dB may be altered; it is known as the *adjusting delay*. There is a rule for adjusting dB depending on the subject's choices: if the subject consistently chooses the small ('fixed', 'unadjusting') reinforcer, the delay to the large reinforcer is reduced, while if the subject prefers the large reinforcer, the adjusting delay is increased. (It is assumed that subjects are sensitive to the changes in the adjusting delay.) The objective is that the adjusting delay tends to an equilibrium value dB', the 'indifference point' at which the effect of the delay of reward B cancels the effect of the larger magnitude of the reward and the two levers are chosen equally often. In practice, trials are usually grouped into blocks of four. The first two trials are forced presentations of each alternative separately, to ensure that the subject samples the currently programmed delays and reinforcers. The other two are free-choice trials. If the subject chooses the same alternative on both of these trials, the delay dB is altered according to the rules stated above. If the subject chooses each alternative once, dB is not altered. Subjects perform this task until dB has reached a stable value (various definitions of stability have been used) and the mean value of dB for stable trials is taken as dB'.

This task has provided strong support for the view that the effects of delayed reward are well described by a hyperbolic discount function (Mazur, 1987), and has been used with success in describing subjects' choice with delayed, probabilistic, and conditioned reinforcement (reviewed by Mazur, 1997). The effects of motivational and neurochemical manipulations have been clarified using this task (Wogar *et al.*, 1992; 1993b) and a version in which the magnitude of the reward is varied according to the same principles has also proved useful (1997a; Richards *et al.*, 1997b; 1999).

So far, this success has been on the 'molar' timescale; that is, based on values of dB' that are the mean of dB over a long series of choices on the part of the subject. The present study was designed to investigate choice behaviour in this task at a 'molecular' (trial-by-trial) level by examining the relationship between dB and choice, and to see if the task was suitable for neurotoxic lesion studies and acute pharmacological studies of impulsive choice. Rats were trained on the adjusting-delay task and their choices analysed to determine their sensitivity to dB. As a simple relationship between dB and choice was not found, computer simulations were conducted to investigate which observed features of performance can be explained by factors independent of dB, and evidence was sought of rats' sensitivity to their history of recent delays.

EXPERIMENT

Methods

Subjects

Eight experimentally naïve male Lister hooded rats were housed in pairs, provided with free access to water and were maintained throughout the experiment at 85% of their free-feeding mass. Housing conditions were described in detail in Chapter 2.

Adjusting-delay technique for assessing choice with delayed reinforcement

This behavioural task was based on those of Mazur (1987; 1988) and Wogar *et al.* (1992; 1993b). Four of the standard operant chambers were used (see Chapter 2), except that they were not fitted with dippers or traylights. The reinforcers used were 45-mg sucrose pellets (Rodent Diet Formula P, Noyes, Lancaster, NH). The apparatus was controlled by software written by R.D. Rogers, N. Daw and R.N. Cardinal.

Rats were first trained to press both levers (FR1 schedule, one-pellet reinforcer) in 30-min sessions daily, until a criterion of 50 presses per session was reached. The two levers were designated Levers A and B, counterbalanced left / right across subjects. Lever A produced immediate small rewards (1 pellet), while Lever B produced delayed larger rewards (2 pellets).

At the start of a session the houselight was switched on, and remained on for the duration of the session. Each session contained 10 trial blocks. Each block consisted of four lever presentations. The first two were forced-choice situations, with Levers A and B presented singly; the A / B order was randomized. Following these, there were two open-choice presentations of both levers simultaneously. Every presentation began with the illumination of the central magazine light, and the levers were extended 10 s later.

When the rat responded on a lever, the light above that lever was switched on, the magazine light was extinguished, and the levers were retracted. When the rat responded on Lever A, one pellet was delivered immediately. When it pressed Lever B, a delay ensued, after which two pellets were delivered. In both cases, the lever light was switched off as pellet delivery commenced. If the rat did not respond on a lever after a 'limited hold' period of 10 s, an omission was scored: the magazine light was switched off and the levers were retracted. No extra presentations were given to make up for omissions, but omissions were a very infrequent event (see *Results*).

If Lever A was chosen on both open-choice presentations, the delay associated with Reward B was decreased by 30% for the next trial block. If Lever B was chosen on both presentations, the delay was increased by 30%. If each lever was chosen once, the delay was not altered. The delay was initially 2 s and was kept within the range 2–20 s; this range was increased to 2–45 s from session 21 (trial block 201) as it became apparent that some rats had reached the maximum delay. From session 64, the delay was altered by 20% rather than 30%.

So that the choice of lever could not affect the frequency of reinforcer delivery, the time between lever presentations was kept constant at 45 s (or 70 s, after the maximum delay was increased). There were 10 trial blocks (of 4 lever presentations) per session, for a session length of 30 min (or 47 min after the increase). Adjusting delays for each subject were carried over from one session to the next as if there were no break.

Subjects were trained on this task for 80 sessions with one session per day.

Analysis of behavioural data

Choice-by-delay graph. To determine whether the current adjusting delay actually influenced the rats' choice, choice-by-delay plots were constructed. To create these plots, omissions were first excluded. Next, subjects' responses were assigned to a bin based on the adjusting delay that was operative at the time the response was made. (As the adjusting delay was altered on a logarithmic scale, bins of 0.1 $\log_{10}$ units were used, though none of the analytical techniques used assumed that time was subjectively perceived by the subjects on a logarithmic scale.) For each rat, in each bin, a preference score was then calculated as the proportion of choices in which the delayed reward was selected.

To supplement this analysis, a simple measure of each rat's sensitivity to delay was derived by calculating the correlation between the rat's choice (Unadjusted lever, scored as 0, or Adjusted lever, scored as 1) and the logarithm of the adjusting delay was calculated. All data from every choice trial (except omission trials) were used, giving up to 1600 data for most rats. As the preference variable was dichotomous, the point-biserial correlation r_{pb} was calculated. This is numerically identical to the Pearson product-moment correlation coefficient r, and may be tested for significance in the same way (Howell, 1997, pp. 257/279–283). Once a correlation coefficient had been computed for each rat, the group's coefficients were compared to zero using a two-tailed t test to establish whether the group exhibited sensitivity to the adjusting delay.

Cross-correlations of preference and adjusting delay. In an attempt to elucidate the causal relationships between the adjusting delay and subjects' preference, cross-correlations were computed. Each rat's complete data set was examined using non-overlapping 'windows' of 10 choice trials (examining choice trials 1–10, then trials 11–20, and so on). Within each window, the preference for the adjusting alternative was calculated as the proportion of choice trials on which the adjusting alternative was chosen. For the same window, the mean $\log_{10}$(adjusting delay) was also computed. The calculated preferences and the mean adjusting delays were placed in temporal order to form two time series, and the cross-correlation function (CCF) of the two time series was computed. (Essentially, a cross-correlation computes the correlation between two functions at different lags and leads.)

This analysis attempts to separate out the influence of preference on delay from the influence of delay on preference, establishing the direction of causality. As preference was programmed to affect the adjusting delay in this task, it was expected that delays would be *positively* correlated with preference scores from the recent past (because preference was scored from 0, being exclusive preference for the unadjusting alternative, to 1, being exclusive preference for the adjusting alternative). Similarly, if long delays were aversive to the subjects, as might be anticipated, it was expected that preference would be *negatively* correlated with delays from the recent past (equivalently, that delays would be negatively correlated with preference in the immediate future).

Mathematical background and pre-processing of data. Cross-correlation, a method within the discipline of time series analysis, depends upon a number of assumptions (see McCleary & Hay, 1980, pp. 229–273; Gottman, 1981, pp. 321–322). The topic is extremely complex and a thorough treatment will not be presented here. However, interpreting a CCF requires that both variables be 'stationary' — loosely, that there be no *autocorrelation* in either variable. (A variable exhibits autocorrelation, or is 'non-stationary', when its value at some time point can be predicted from the value of the same variable at a different time; a variable that does not exhibit this property is said to be stationary, or 'white noise'.) Autocorrelation in either variable can introduce spurious correlation into the CCF; thus, a cross-correlation of autocorrelated variables is uninterpretable (McCleary & Hay, 1980, pp. 243–246). To correct for this, transformations are conducted before cross-correlating; this process is called 'prewhitening' and is performed on each variable separately, termed 'double prewhitening' (see also Hare, 1996, chapter 1). An example of this technique is given by Bautista *et al.* (1992).

To prewhiten a time series, an autoregressive integrated moving average (ARIMA) technique was used (Box & Jenkins, 1970; McCleary & Hay, 1980, p. 18; Gottman, 1981; see also StatSoft, 1999). Again, this will not be described thoroughly here, but the essence is to build a mathematical model of a time series that describes the autocorrelation in the time series, then to subtract the model's predictions from the original data, removing the autocorrelation from the time series. Briefly, the notation 'ARIMA(p,d,q)' describes a mathematical model of a time series, specifying the degrees to which a time-lagged value of the variable is used as a predictor (autoregression; p), the number of passes on which the variable should be subtracted from a time-lagged version of itself before being used as a predictor (differencing; d), and the number of moving average parameters (q). As an example, an ARIMA(2,1,0) model contains two autoregressive parameters and no moving average parameters, calculated after the series has been differenced once. An autocorrelation function (ACF), which correlates a function with a time-shifted version of itself, may be used to identify the ARIMA model likely to provide the best fit to the data in question. The autocorrelation functions of ARIMA models are characterized by a discrete number of spikes corresponding to the moving average part of the model, and damped exponentials and / or damped sine waves corresponding to the autoregressive part of the model. Autocorrelation and partial autocorrelation functions (ACF, PACF) were computed for each variable being prewhitened (i.e. the preference score time series and the adjusting-delay time series) and used to identify autoregressive and / or moving average terms (that is, particular values of p, d, and q) as described by McCleary & Hay (1980), minimizing the number of terms included in the model (though model parsimony was considered secondary to obtaining a good fit). This model was then fitted to the variable, checking that it provided a significant fit, and the *residuals* were examined. If the residuals exhibited no autocorrelation (were of a white noise type), then the objective had been achieved: the autocorrelation had been removed from the original variable, and those residuals were used for cross-correlation. This technique is sometimes referred to as 'filtering' the original time series through an ARIMA model. It should be noted that the process of fitting an ARIMA model is empirical; a model is fitted to each time series separately (there are two time series from each subject), with the sole objective of removing autocorrelation from that time series. Importantly, the preference score and adjusting-delay time series were prewhitened *independently* before cross-correlation.

Finally, as the usual assumptions of correlation also apply to cross-correlation, the variables entered into the cross-correlation were checked for normality using the Kolmogorov–Smirnov test and by inspection of Q–Q plots (which plot the quantiles of the variable against the quantiles of a normal distribution).

To summarize, the following steps were conducted for each subject:

1. Calculate windowed choice ratios and log(adjusting delay), to give two time series.
2. Generate and fit an appropriate ARIMA model to each time series.
3. Cross-correlate the residuals from the two fitted ARIMA models.

Window size. The results of cross-correlational analysis depend in part upon the 'window' size used — for example, large windows permit more accurate calculation of preference, but they also obscure rapid, high-frequency changes in the cross-correlation coefficient. Pilot analyses were conducted with window sizes of 5, 10, 20, and 40. Smaller window sizes were not used, to avoid the preference score approaching a dichotomy, which would have violated the assumptions of the analysis. In all cases, the maximally *significant* cross-correlations were observed with the minimum window size used (5 choice trials); that is, the 'optimal' window size for detecting a correlation did not vary across subjects. The prewhitened data subjected to cross-correlation approximated a normal distribution even with this small window. Furthermore, the use of windows larger than 5 did not, in general, alter the lag at which the maximum cross-correlation was observed. As would be expected, larger windows yielded larger numerical correlation coefficients, but also increased the width of the confidence interval (as a larger window reduces the number of windows being analysed). Therefore, a window size of 5 was used for all subsequent analyses. Cross-correlations were computed out to lags and leads of 200 choice trials (40 decision windows).

Results

One rat (subject C7) fell ill and ceased responding from session 72; subsequent data from this rat were discarded. Other than this, responding was reliable, with rats failing to press a lever on only 1.53% of presentations. The obtained adjusting delays for 80 sessions (800 trial blocks) are shown in Figure 56 and typical individual records are shown in Figure 57. It is apparent that although the mean of the group of subjects appears relatively stable in the range 10–15 s, values of dB for individual subjects varied widely across the permissible range (which was 2–20 s for the first 20 sessions, and 2–45 s for the remainder).

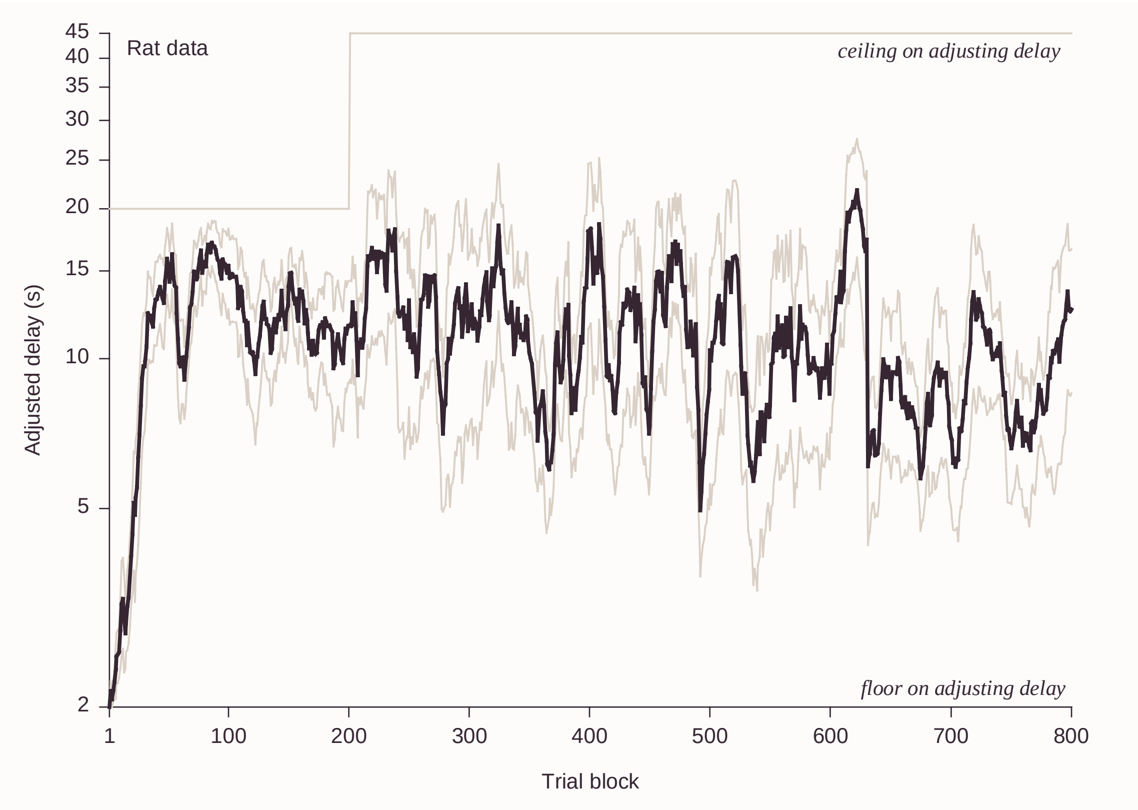


Figure 56. Group mean adjusting delay for 8 rats, displayed by trial block. Thick and thin lines show mean ±1 SEM. The boundaries between sessions are not shown. The maximum permissible value of the adjusting delay is shown as a stepped line at the top of the figure; this maximum was increased after session 20 (trial block 200).

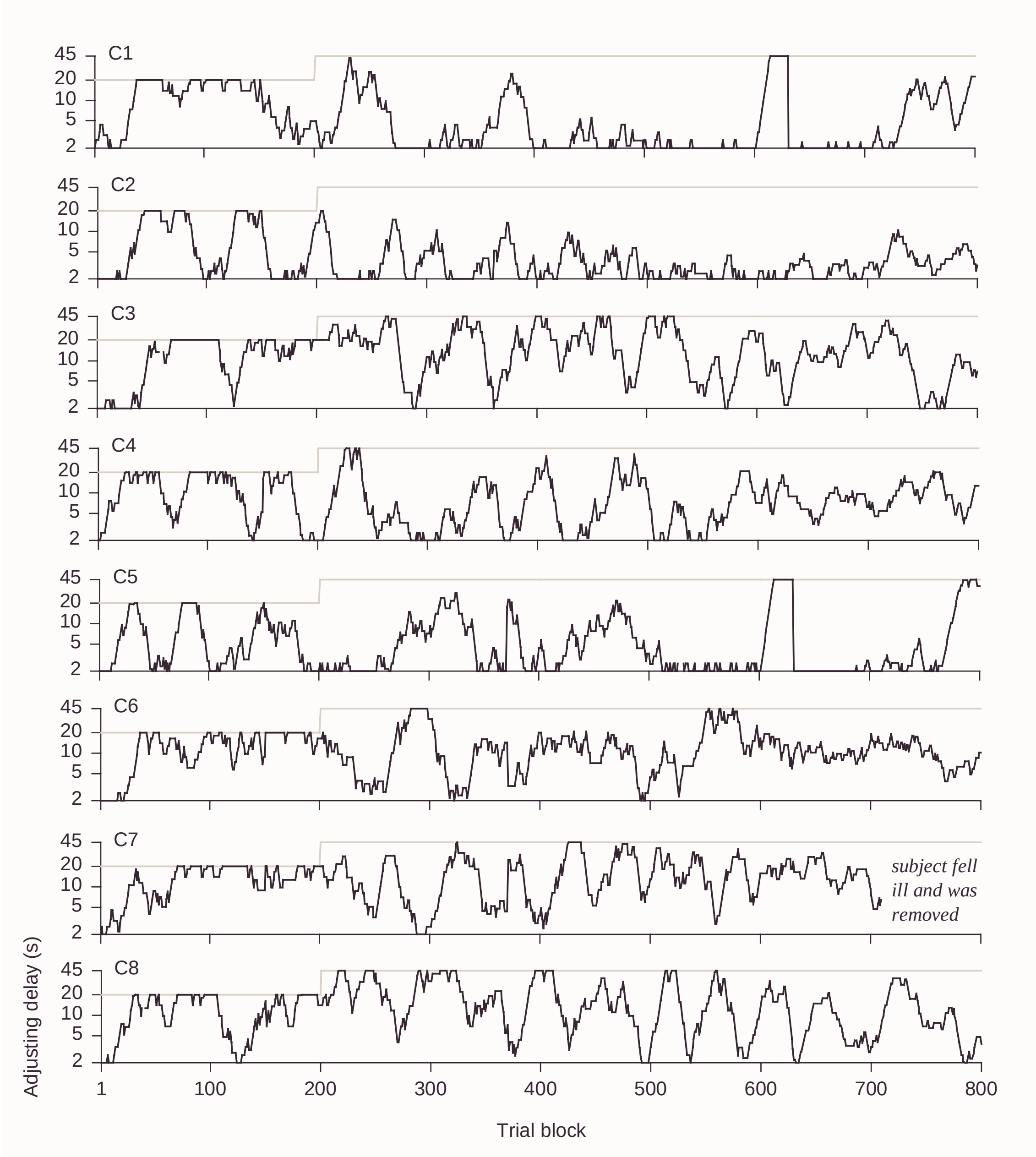


Figure 57. Individual records, for all trial blocks. The thin grey line shows the maximum permissible value of the adjusting delay, as in Figure 56.

Choice-by-delay plots. Choice-by-delay plots are shown in Figure 58. As preference scores were arbitrarily calculated such that 0 represents exclusive preference for the unadjusted alternative (lever A) and 1 represents exclusive preference for the adjusted alternative (lever B), the theoretically predicted result would be a line of negative slope, indicating reduced preference for the large reinforcer at long delays. The obtained curve is relatively flat, indicating no effect of delay. If anything, Figure 58 suggests that a number of subjects had *high* preferences for the delayed reward when the delay was longest, and low preferences when the delay was low. A plausible interpretation is that the rats had a tendency to repeat their last response at extremes of delay — for example, a subject pressing the unadjusted lever many times in succession will drive dB down to its minimum permissible value, after which the subject can accumulate 'unadjusted' responses at the minimum delay. As a group, the point-biserial correlations did not differ from zero (t_7 = 1.599, NS; these data are also plotted later in Figure 61C, p. 161).

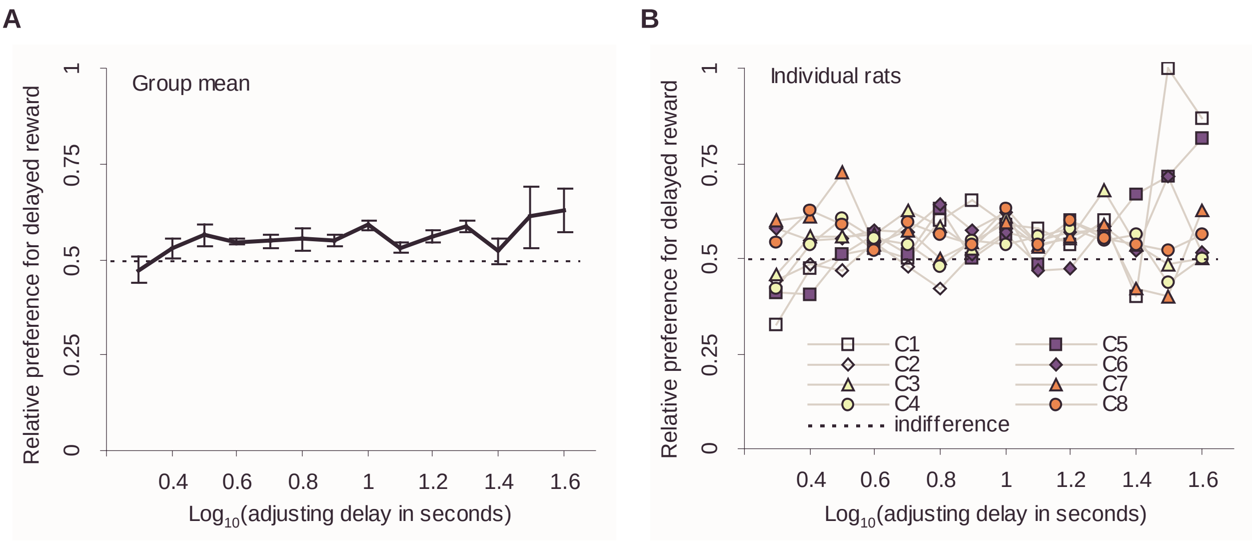


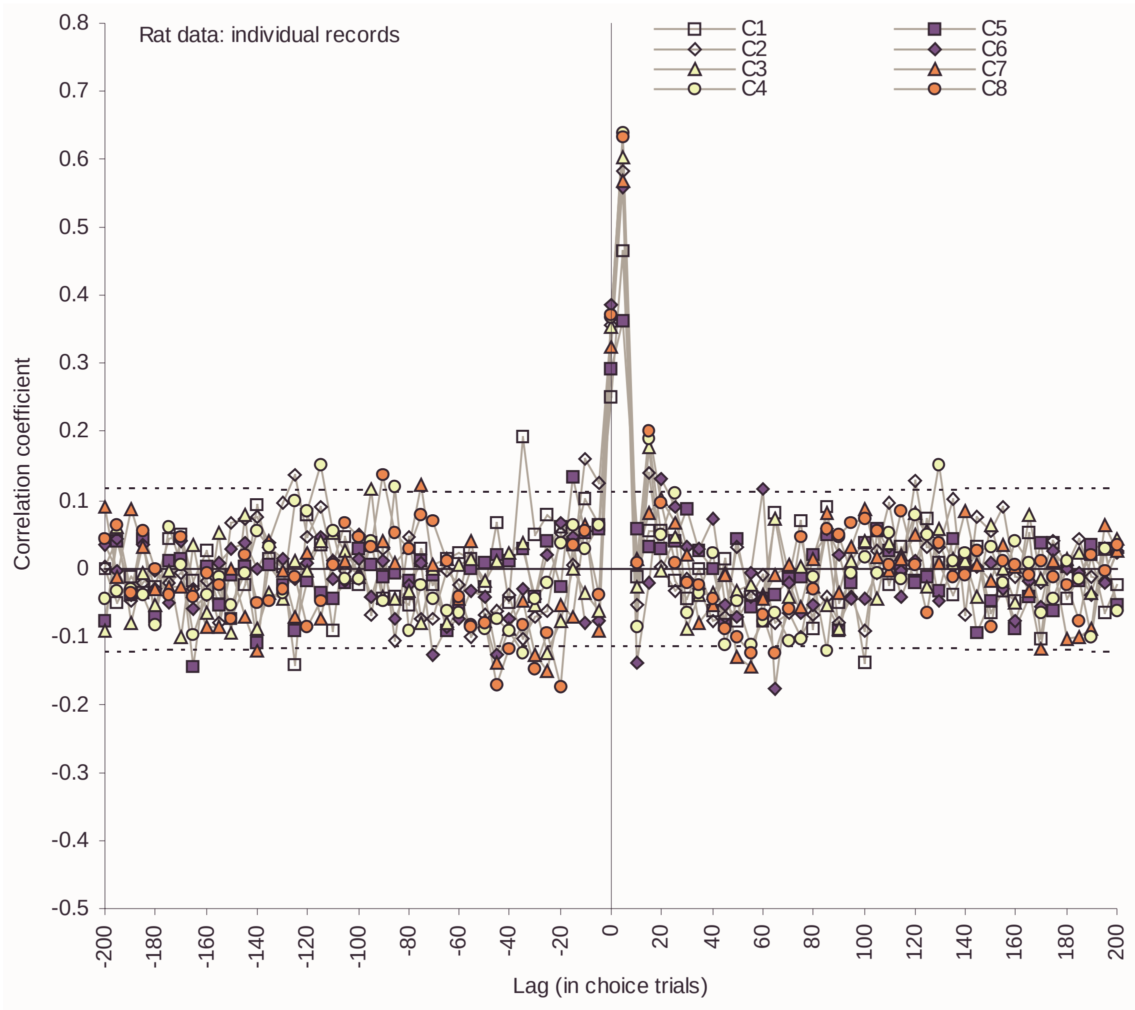
Figure 58. Choice-by-delay graphs for 8 rats. The ordinate (vertical axis) represents a preference score, from 0 (exclusive choice of the unadjusted, immediate lever) to 1 (exclusive choice of the adjusted, delayed lever), with omissions not analysed. The abscissa (horizontal axis) is $\log_{10}$(adjusting delay); preference was calculated in bins of 0.1 log units. **A:** Group mean ± SEM. **B:** Individual subjects.

Slow changes in preference? It is probably unreasonable to expect rats to be perfectly sensitive to the adjusting delay currently in force. An obvious alternative is that the subjects are not immediately sensitive to changes in dB, despite the forced-choice trials, but rely on a slow cumulative learning process that gradually alters preference once the adjusting delay has been suboptimal for some time, leading to 'overshooting' and oscillation. (For example, a subject might prefer the large, delayed reinforcer when dB is low, leading to an increase in dB, yet fail to adjust its preference to reflect that increase for some time. By then, dB would have increased beyond the subject's point of indifference, the small reinforcer would be preferred and the cycle would reverse. The value of dB would therefore oscillate around the indifference point rather than converging to it.) This is the view of several investigators (C.M. Bradshaw, personal communication, 7 October 1998; J.E. Mazur, personal communication, 16 November 1998). It may be termed a 'running average' hypothesis, since it suggests that the subjects are sensitive to some form of average of several recent values of dB.

If the 'running average' hypothesis is correct, then it is not so surprising that the choice-by-delay curve might be flat in its middle region. If dB oscillates around in the indifference point, there will be a range of values of dB for which the subject sometimes chooses the unadjusted alternative (at times when it is driving dB down), but sometimes chooses the adjusted alternative (when it is driving dB up). These tendencies might cancel out, leading to apparent indifference for this range of values of dB.

Cross-correlelograms. However, the 'running average' hypothesis predicts that choices should be correlated with adjusting delays from the recent past. Consequently, preference scores were cross-correlated with adjusting delays (as described in the *Methods*). For the prewhitening phase, it was found that an ARIMA(1,0,0) consistently described the vast majority of autocorrelation in the choice ratio time series, and an ARIMA(1,0,1) model was used for the delay time series. The final cross-correlations are plotted for each rat in Figure 59.

A



B

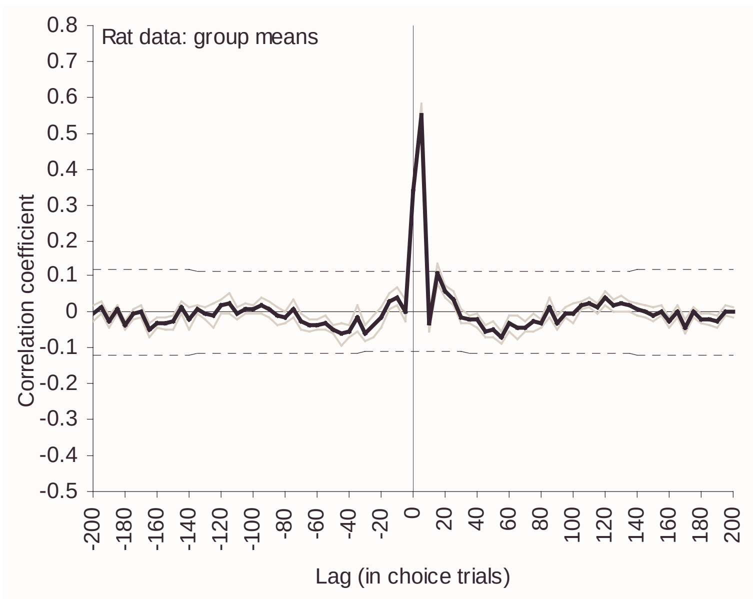


Figure 59. Cross-correlation of preference for the adjusting alternative with the adjusting delay. The analysis is organized so that positive lags indicate the effect of choice on delay, and negative lags indicate the effect of delay on choice. Thus, the ubiquitous positive correlation at small positive lags is the programmed rule for adjusting delay: when preference for the adjusting alternative is high, the delay is increased for later trials. A negative correlation at negative lags would indicate that high delays reduce subjects' subsequent preference for the delayed alternative. Confidence limits (horizontal dotted lines) are 2 SE. **A:** individual rats. **B:** group mean ± SEM.

Figure 59 clearly reveals the contingencies programmed into the task: that choice affects delay, such that high preference for the adjusting alternative is strongly correlated with the adjusting delay in the near future. In contrast, it is not clear at all that the adjusting delay affected choice behaviour. (If the delay did affect choice in the theoretically sensible direction, negative cross-correlations would be expected at negative lags in Figure 59.) Subjects' CCFs did exhibit occasional peaks in this region; to estimate the time course of the subjects' apparent sensitivity to the adjusting delay, the largest cross-correlation coefficients for each rat are printed in Table 15. It can be seen that there is considerable variation in the lag at which subjects' preferences were maximally correlated with the adjusting delay; some subjects' preferences appeared to be affected by the average adjusting delay from 15–40 choice trials previously (equivalent to 8–20 trial blocks, or 1–2 sessions), some subjects showed the maximal peak at 80–120 choice trials (or up to 6 sessions) previously. Additionally, none of these peaks is large. If the group is considered as a whole (Figure 59B), it is clear that no consistent sensitivity to past delays is seen.

Table 15. Maximum cross-correlations for each subject. The cross-correlations were computed using windows of 5 choice trials, and the correlation coefficient that was largest relative to its standard error at a negative lag is listed, as an index of the effect of delay upon the subject's choice. No attempt has been made to correct for multiple comparisons.

Rat	Lag (choice trials)	Correlation coefficient	Confidence limit (2 × standard error)
C1	–125	–0.141	–0.116
C2	–85	–0.107	–0.114
C3	–25	–0.123	–0.112
C4	–35	–0.125	–0.114
C5	–165	–0.146	–0.118
C6	–45	–0.127	–0.114
C7	–25	–0.15	–0.112
C8	–20	–0.173	–0.112

COMPUTER SIMULATIONS

It is intriguing that subjects performing a task that has produced highly consistent end-points in other studies (see Mazur, 1987; 1988) should show apparent insensitivity to dB. To establish what performance is possible using a decision rule that does not take account of the delay to reinforcement, a number of decision rules were simulated on a computer.

Methods

Computer simulations were written in the programming language C++ (Stroustrup, 1986; 2000); data from the simulations were fed into the same means of analysis as those from the real rats. Six decision rules were simulated, as follows:

Random. Decisions under the Random rule were independent of the adjusting delay. The adjusting alternative was chosen with probability 0.5 (and the unadjusting alternative also with probability 0.5).

Biased. The Biased rule was also delay-independent. The overall frequencies with which each rat chose the two alternatives were calculated (ignoring trials on which an omission occurred); each simulated subject was assigned the relative preference of one of the rats as its bias. On each choice trial, the adjusting alternative was selected with that probability, as shown in Table 16.

Biased-60. The Biased-60 rule implemented a fixed bias; the adjusting alternative was chosen with probability 0.6 (and the unadjusting alternative with probability 0.4).

Markov Chain. A Markov chain is an abstract entity that can be in one of several states at any given moment (phrased more obscurely, a 'finite state machine'). The chain is characterized by the set of probabilities of a transition occurring between each possible pair of states. In the present task, each choice alternative can be represented as a state (Adjusted and Unadjusted). A transition from the Adjusted state to the Unadjusted state would then represent a rat choosing the Adjusted lever on one choice trial, and the Unadjusted lever on the next trial.

Transition probabilities were calculated for each rat. After discarding omission trials, all choice trials were placed in order, and the relative frequency of the four possible transitions were computed. The transition probabilities are shown in Table 16. Eight Markov chain simulations were then performed, each simulation having the characteristic transition probabilities of one of the rats. The first choice made by each simulation was also the same as that of its corresponding rat.

Table 16. Overall proportion of choice responses on which the adjusting (Adj) alternative was chosen (used for the Biased rule), together with transition probabilities for each rat (used for the Markov chain decision rule). As omissions were ignored, pairs of transition probabilities sum to 1. The final column shows the first choice response ever made by each rat.

Rat	Overall proportion of Adj responses	*p*(Adj → Adj)	*p*(Adj → Unadj)	*p*(Unadj→ Adj)	*p*(Unadj→ Unadj)	First response
C1	0.470	0.544	0.456	0.406	0.594	Adj
C2	0.478	0.495	0.505	0.465	0.535	Adj
C3	0.570	0.630	0.370	0.496	0.504	Unadj
C4	0.530	0.579	0.421	0.476	0.524	Unadj
C5	0.489	0.573	0.427	0.409	0.591	Adj
C6	0.557	0.557	0.443	0.557	0.443	Adj
C7	0.551	0.554	0.446	0.546	0.454	Adj
C8	0.565	0.590	0.410	0.534	0.466	Unadj

Preference. The Preference rule was intended to mimic a ‘perfect’ subject — one whose choices immediately and accurately reflected the programmed adjusting delay. Each subject was assigned a preferred delay. On each choice trial, if the adjusting delay exceeded the preferred delay, the unadjusting alternative was chosen. If the adjusting delay was lower than the preferred delay, the adjusting alternative was chosen. If the adjusting delay exactly matched the preferred delay, the subject chose randomly ($p = 0.5$ for each alternative).

In order to match the Preference rule to the rats, each simulated subject was assigned a preferred delay derived from data from one rat; this preferred value was taken to be the mean adjusting delay over the last 200 trial blocks of testing (blocks 601–800). These values, in seconds, were 9.85 (subject C1), 3.55 (C2), 12.43 (C3), 9.20 (C4), 11.16 (C5), 10.83 (C6), 11.44 (C7: as this subject fell ill, its mean was calculated from trial blocks 601–710 only), and 11.50 (C8). The mean preferred delay for the simulations was thus 10.0 ± 0.99 s.

Running Average. The Running Average rule was also delay-dependent; it used the same basic decision rule as the Preference rule, and the preferred delays were calculated in the same manner. However, instead of comparing its preferred delay to the adjusting delay operative at that moment, the Running Average rule compared its preferred delay to the mean of dB over the last several choice trials. The actual decision window varied from subject to subject, and were chosen arbitrarily. The decision window sizes were drawn from a normally distributed random variable with a mean of 40 choice trials and an SD of 20 choice trials. The actual values used were 10, 60, 55, 52, 40, 22, 70, and 35 choice trials (mean 43, SEM 7).

It should be noted that the simulated Running Average rule has high mnemonic demands — the subjects remember every single delay within their decision window — and biologically more plausible algorithms exist (see Killeen, 1981, for a discussion), but it is a simple illustration of sensitivity to past delays that does not give heavy weighting to the most recent value. As more plausible algorithms often *do* give heavier weighting to more recent values (e.g. exponentially-weighted moving average; Killeen, 1981), the Running Average rule represents a stringent test of the analytical technique of cross-correlation as applied to this situation — if cross-correlation is observed with this rule, it would certainly be expected with more plausible algorithms.

Table 17 summarizes these decision rules.

Table 17. Summary of simulated decision rules. (Adj = selection of the adjusting alternative; Unadj = selection of the fixed alternative.)

Simulation name	Choice rules
Delay-independent rules	
Random	p(Adj) = 0.5; p(Unadj) = 0.5
Biased	The probability of selecting each alternative was fixed, and set to the overall probability with which one of the rats chose the alternatives (see Table 16).
Biased-60	p(Adj) = 0.6; p(Unadj) = 0.4
Markov chain	The probability of choosing each alternative was based solely on the previous choice, with the transition probabilities shown in Table 16.
Delay-dependent rules	
Preference	• if delay < preference, p(Adj) = 1; p(Unadj) = 0 • if delay > preference, p(Adj) = 0; p(Unadj) = 1 • if delay = preference, p(Adj) = 0.5; p(Unadj) = 0.5 Each subject had its own preferred delay, matched to one rat; these delays had a mean of 10.0 ± 0.99 s.
Running Average	Choice is determined as for the Preference rule, but the delay used to make the decision was the mean of dB over the 43 ± 7 most recent choice trials (see text).

For all decision rules, the starting conditions and the rules for updating the adjusting delay based on the subject's choice were identical to those in the real task (described earlier), including the change in the limits set on dB. Six decision rules were simulated, with 8 simulated subjects in each condition. Simulations were not repeated.

Application of a stability criterion to the Random decision rule. In a separate simulation, the Random rule was also used to establish the length of time needed for a randomly-deciding subject to meet the stability criteria previously used for pigeons by Mazur (1987; 1988; personal communication, 22 October 1998). The task simulated was changed so it matched exactly that used by Mazur (1988), as the starting value and stability criteria were unspecified in Mazur (1987). Thus, the starting adjusting delay was 8 s; trials were grouped into blocks of two single-lever trials and two choice trials, as before; the adjusting delay was altered arithmetically in steps of ±1 s; the minimum value of dB was 0 s, and there was no maximum set on dB (as the time between reinforcement and the next trial was held constant in Mazur's study, rather than the time between the start of two consecutive trials). There were 64 trials per session (32 choice trials), the adjusting delay from one session was carried over directly into the next session, and subjects were tested for a minimum of 10 sessions. Data from the first two sessions were discarded and the rest of the data were tested for stability as follows. Each session was divided into two 32-trial blocks (i.e. 16 choice trials) and the mean adjusting delay for each block was calculated. The stability criteria were: (1) that neither the highest or the lowest single-block mean could occur in the last six blocks; (2) that the mean adjusting delay across the last six blocks was not the lowest or the highest such six-block mean; (3) that the mean of the last six blocks was within 10% of the mean of the preceding six (or within 1 s, whichever was greater). One hundred instances of the Random rule were simulated, and the time taken for each to meet these stability criteria was recorded.

Relationship between bias and mean adjusting delay. For a subject that takes no account of the adjusting delay, it is likely that the subject's bias has a systematic effect on the obtained mean adjusting delay, dB'. Firstly, to establish whether a manipulation that influenced a subject's bias could in principle affect dB', the mean value of dB was calculated over trial blocks 400–800 for each simulated subject using the Random or Biased-60 decision rules, yielding one value of dB' per subject (and n = 8 per group); these values were then subjected to a univariate ANOVA with the decision rule as a between-subjects factor.

Secondly, to establish the quantitative nature of the relationship between bias and dB’, simulations were conducted using the conditions of Mazur (1988) described above. Each simulated subject was assigned a bias towards the adjusting lever; on every choice trial, it chose the adjusting lever with $p(\text{Adj}) = bias$, and $p(\text{Unadj}) = 1 - bias$. At every level of bias from 0.4 to 0.6 in steps of 0.01, one hundred subjects were simulated. The stability criteria described above were applied, and the mean value of dB over the last six (stable) half-session blocks was measured, just as in Mazur (1988, Table 1). This simulation was also repeated with dB limited within a range of 0–40 s.

Results

Local analysis of the simulated decision rules

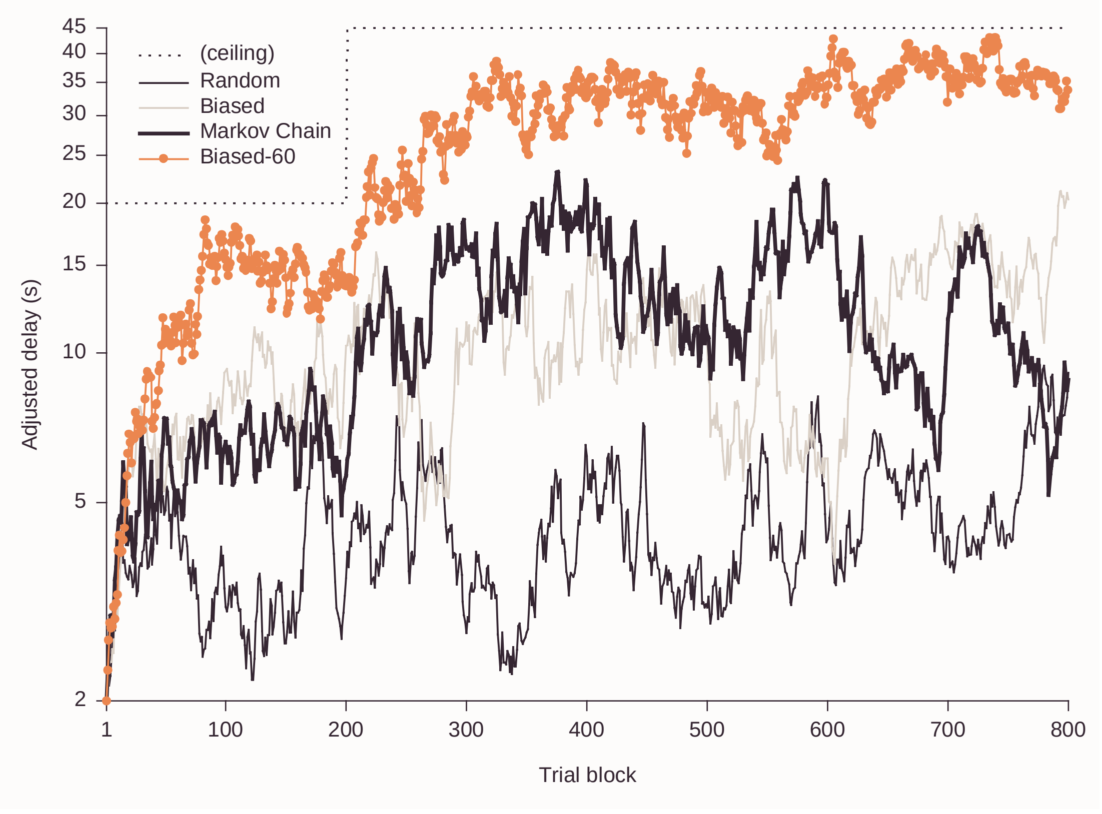
The evolution of the adjusting delay is shown for the simulated decision rules in Figure 60 (compare the rat data in Figure 56). The Random rule simply generates a random walk between the limits set on dB. The Biased-60 rule chooses the adjusting alternative more frequently than the fixed alternative and thus drives the adjusting delay to high values, while wide excursions in dB are seen in the Random rule. The Preference rule generates tight oscillations around the preferred delay; even though the simulations’ preferred delays were taken from the rat data, this simulation generates much less variability than the rats. The Running Average rule produces a sinusoidal oscillation around the preferred delay. As the group means shown in Figure 60 were derived from eight simulations, each with a different decision timebase, the group mean is not perfectly periodic (in fact, it represents a spectrum with eight frequency components). The delay-independent rules produced the pattern most like the rat data, with the Biased and Markov Chain rules generating values of dB in a similar range to the rats.

Figure 61 shows choice-by-delay plots for the simulated decision rules. This form of plot is clearly inadequate to demonstrate all but the simplest form of delay sensitivity: only the simple Preference rule demonstrates the theoretically predicted curve (high relative preference for the adjusting alternatives at low delays, and low preference at high delays). The other curves, including the delay-sensitive Running Average rule, are essentially flat. Indeed, the Running Average rule shows a *reduced* preference for the adjusting alternative at the minimum delay, probably due to repetition of responses — when the delay is high, for example, this rule begins to choose the Unadjusted lever and drives the delay to the minimum value; however, as its decisions are based upon several recent delays, it does not ‘notice’ that the delay has reduced for several trials, during which time it accumulates Unadjusted responses at the minimum delay.

Cross-correlational analyses of the decision rules were then conducted. The prewhitening process will be described first.

Prewhitening. For the Random, Biased, and Biased-60 simulations, as would be expected, there was no autocorrelation in the choice ratios but there was autocorrelation in the delays. As the adjusting-delay task only permits values of dB that are the same as, or a small way from, the value of dB in the preceding trial block, the delay at time t is correlated strongly with the delay at time $t - 1$; thus, this autocorrelation was modelled successfully by an ARIMA(0,1,0) model.

The same was true of the Markov Chain simulation. Even though each choice was programmed to depend (to a small extent) on the previous choice, the autocorrelation in choice ratios did not reach significance and no correction was made for it. The delay autocorrelation was again described by an ARIMA(0,1,0) model.



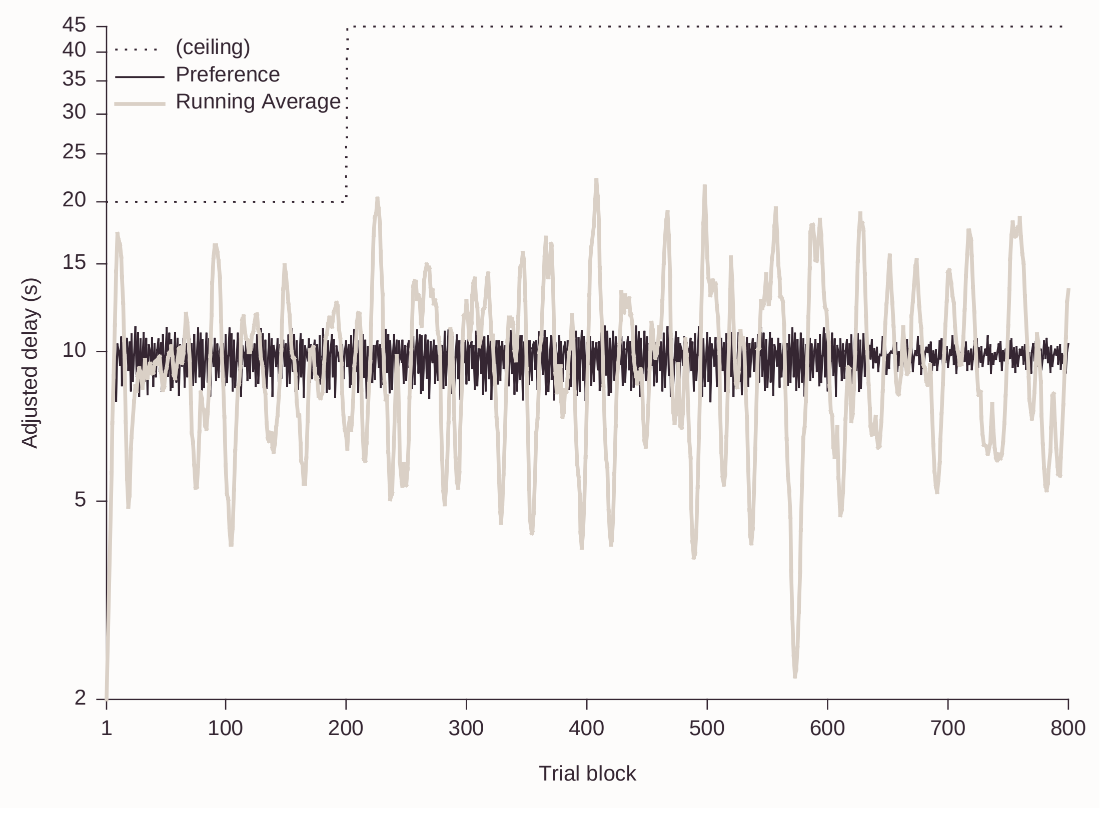


Figure 60. Group mean adjusting delay for the simulated decision rules, displayed by trial block. All simulations have $n = 8$. The top panel shows the delay-independent decision rules, and the bottom panel shows the delay-dependent rules.

The Preference simulation exhibited a significant autocorrelation of choice ratios (because its decisions at any one moment are closely related to decisions in the recent past, as the rule oscillates about its preferred value). This was removed by filtering the choice ratio time series through an ARIMA(2,0,1)

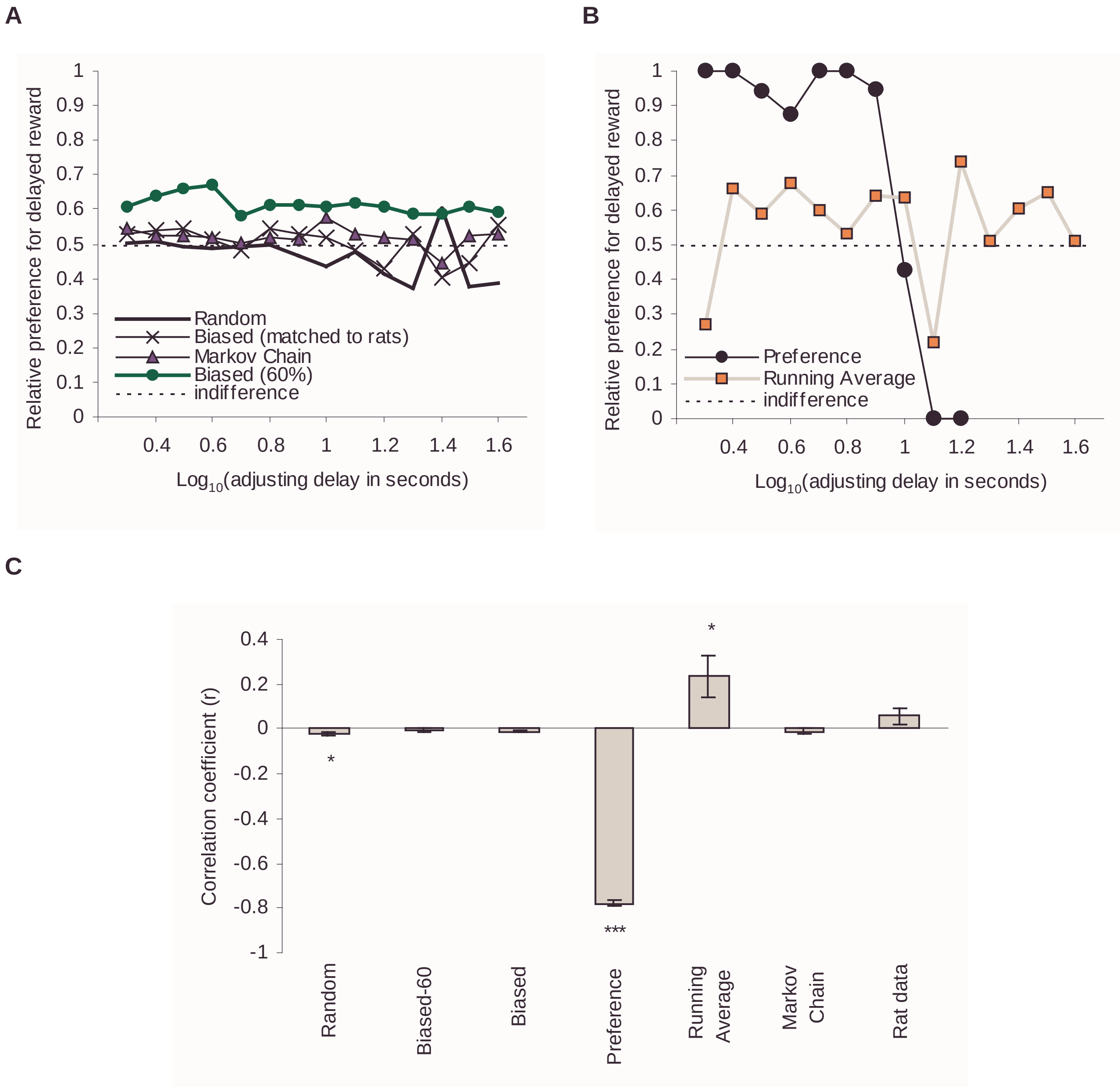


Figure 61. A & B: Choice-by-delay plots for the simulated decision rules. Only the simple Preference rule clearly demonstrates delay sensitivity in this plot, even though the Running Average rule is also delay-sensitive. **C:** Mean (± SEM) correlation coefficients for the correlation between preference for the adjusting lever and the adjusting delay. Correlation coefficients were calculated for each rat or simulated subject using data from all that individual's choice trials; the correlation coefficients were then compared to zero as a group using a two-tailed *t*-test (* $p < .05$; *** $p < .001$). The Preference rule exhibits consistent, negative correlation between preference and delay, indicating that it chooses the adjusting lever when the delay is low, and vice versa. The Running Average rule, which chooses on the basis of delays from the recent past, exhibits a small *positive* correlation between preference and the delay that is operating at the actual moment of choice. Ironically, the Random rule exhibits a significant (though very small) negative correlation! No other decision rule exhibited significant correlation; neither did the rats' choices.

model (determined following the method of Gottman, 1981, p. 262), and the delay time series through an ARIMA(4,0,3) model.

By its nature, the Running Average simulation makes decisions that are strongly correlated with decisions from the recent past. The 'recent' past in this case was quite long — the simulation with the 'longest memory' took account of delays from the last 70 choice trials (14 windows of 5 choice trials). Thus, a high-order ARIMA was necessary to capture the autocorrelation: an ARIMA(14,0,0) model was found to remove the vast majority of autocorrelation from both time series.

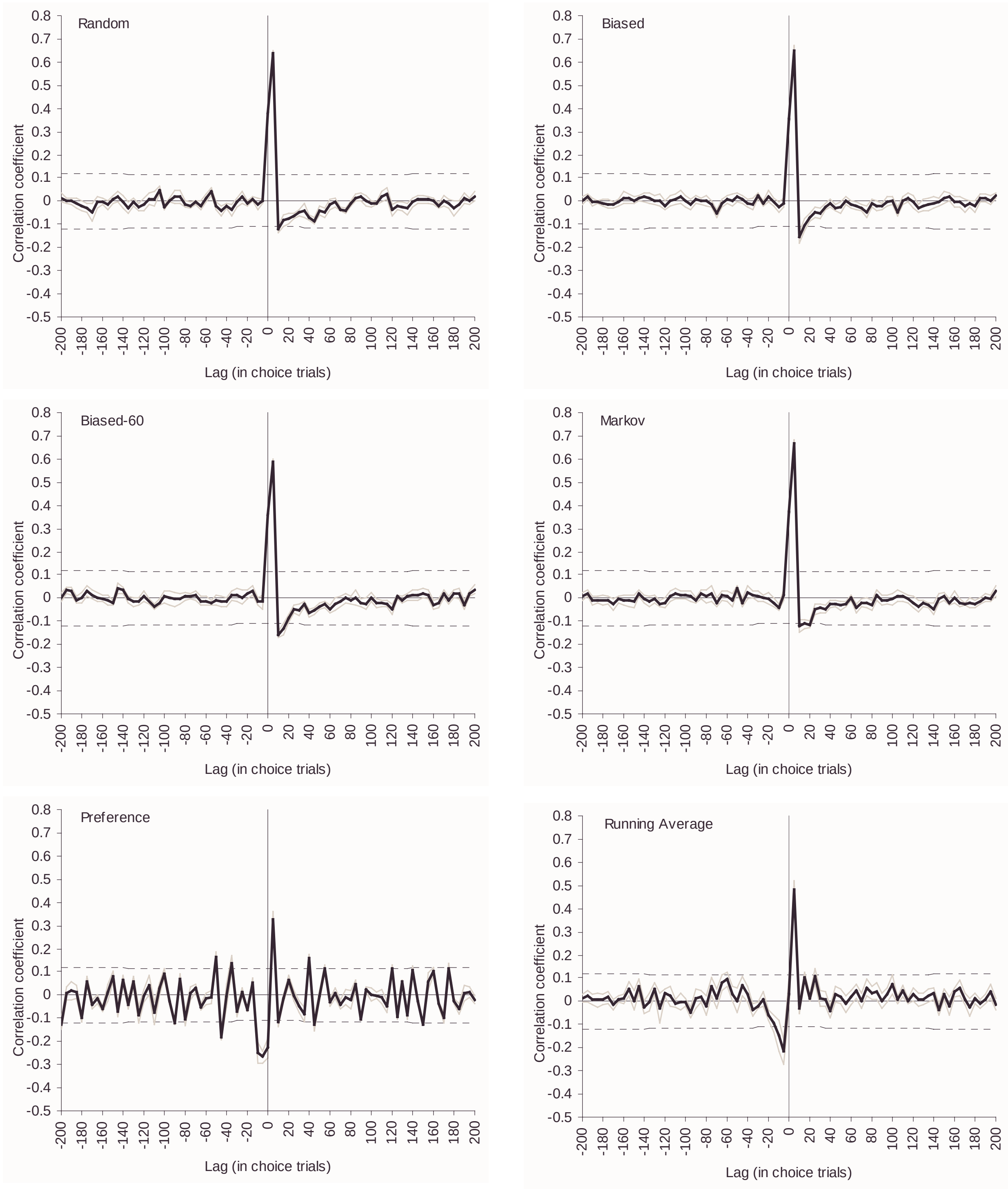


Figure 62. Cross-correlation functions for the simulated decision rules (group means ± SEM), as in Figure 59. Confidence limits (horizontal dotted lines) are 2 SE. All decision rules exhibit a positive cross-correlation at positive lag, a result of the ubiquitous rule through which subjects' choices affect the adjusting delay. Only two rules (Preference and Running Average) exhibit a negative cross-correlation at negative lags, a phenomenon that suggests that the adjusting delay affects the subjects' choices, as indeed was the case for these and only these two rules.

CCFs. The cross-correlation functions for each decision rule are shown in Figure 62. Again, the CCF technique successfully detected the contingencies built into the task (the causal relationship: preference → delay) in all cases. In addition, this technique successfully discriminated between rules that based their decisions upon the adjusting delay, and those that did not. Significant negative cross-correlations at negative lags (suggesting the causal chain: delay → preference) were detected for the Preference and Running Average rules, but for none of the delay-independent rules.

Inspection of individual records of the Random rule (Figure 63) revealed occasional 'significant' negative cross-correlations. As this decision rule was not influenced by the adjusting delay, there are two possible explanations. The first is failure of the prewhitening process to capture all of the autocorrelation in the delay time series; although prewhitening dramatically reduced the degree of autocorrelation, very small autocorrelations occasionally remained, having not been described by the ARIMA model. Autocorrelation can introduce spurious correlation into a CCF (McCleary & Hay, 1980). The second is simple statistical variation. The confidence intervals calculated by the statistical software used take into account the number of data points used to calculate the CCF, but not the number of leads and lags over which the CCF is computed and the number of comparisons this implies. The occasional isolated 'significant' correlation may therefore reflect Type I error (false rejection of the null hypothesis).

The relevance of this discussion is in the comparison with Figure 59A (p. 155), the cross-correlation data for the rats. It may be that occasional negative correlations observed in the rat data are due to the same processes that contributed to correlations in Figure 63, the Random simulations. While it remains possible that the rats exhibited genuine sensitivity to delay, though very slight and with a great deal of variation in its timescale across rats, the rats exhibited no evidence of systematic, consistent sensitivity to the adjusting delay, which is best assessed by consideration of the group mean (Figure 59B).

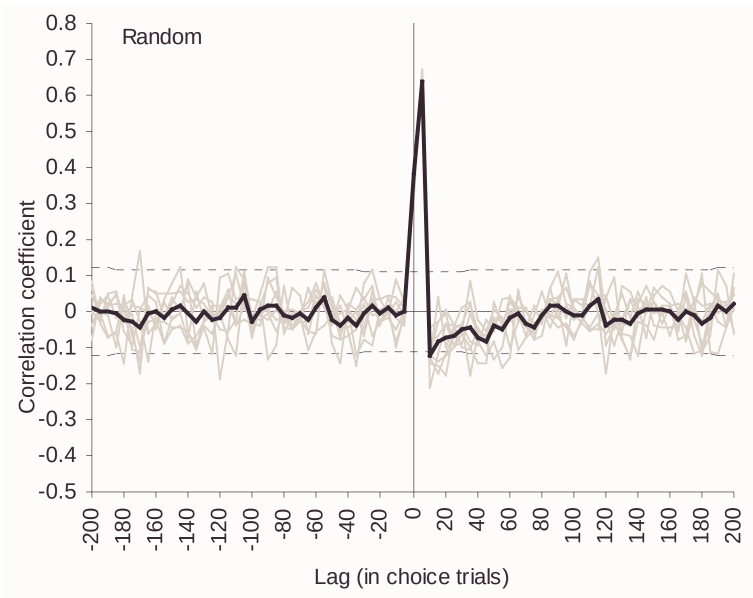


Figure 63. Cross-correlation functions for the simulated Random rule, plotted for each individual simulated subject (thin grey lines), together with the group mean (thick black line). Compare Figure 59.

Achievement of stability criteria by a delay-independent decision rule

Under the task conditions and stability criteria used by Mazur (1988), randomly-deciding simulated subjects reached stability after a mean of 15 sessions (range 10–43, SD 6, with 10 being the minimum number of sessions permitted by the criteria).

For comparison, Mazur (1988) found that pigeons reached stability in a mean of 14 sessions (range 10–29, SD 4; data taken Table 1 of Mazur, 1988, using all 61 conditions experienced by the four pigeons in which two choice trials were given per trial block). Figure 64 illustrates the point at which pigeons in Mazur's (1987) experiment were considered stable.

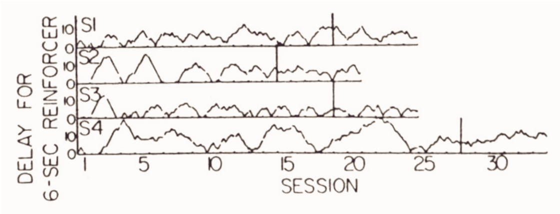


Figure 64. Sample individual dB records of four pigeons, from Mazur (1987). Vertical bars mark the point beyond which performance was considered stable by Mazur's criteria.

Effect of bias on dB' using a delay-independent decision rule

When quasi-stable values of dB' from the Random and Biased-60 decision rules were compared (see *Methods*), it was found that the Biased-60 rule led to significantly higher values of dB' than the Random rule ($F_{1,14} = 649$, $p < 0.001$).

Results of the simulations designed to establish the relationship between bias and dB' are shown in Figure 65. When no limits were placed on dB, biasing the simulated subjects towards the adjusting lever increased the quasi-stable value of dB' (Figure 65A / C), the number of sessions to meet the stability criteria (Figure 65B), and the variance of these two measures. When the maximum value of dB was limited, manipulations of bias produced a sigmoid change in dB' (Figure 65D) with only minor effects on the number of sessions to criterion, which followed an inverted-U-shaped curve (Figure 65E).

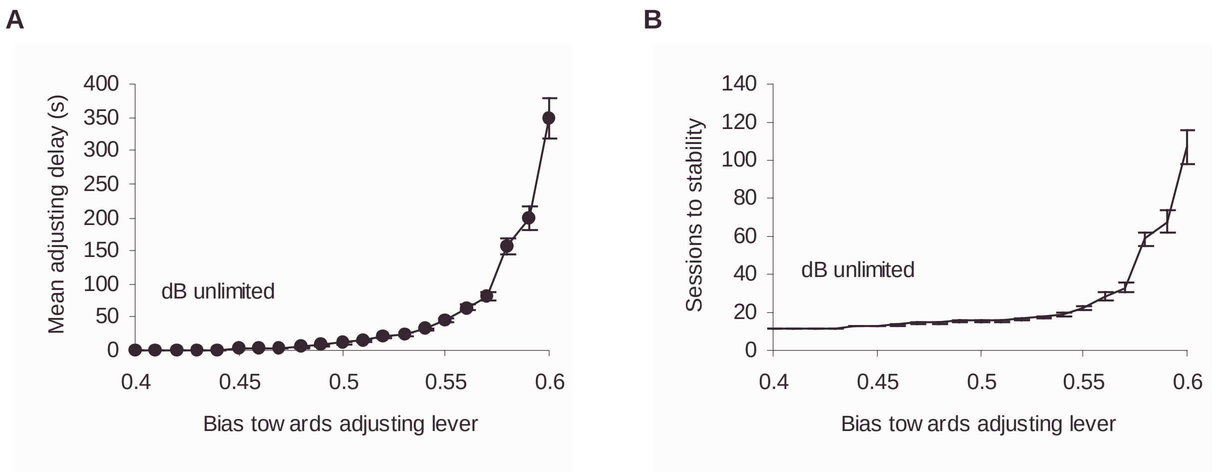
There are regions of the curves in Figure 65A (magnified in Figure 65C) and Figure 65D that are approximately linear. Thus, if a subject chooses between the two levers in a way that is independent of dB, and a manipulation — such as a change in the reinforcement available on the unadjusting lever — were to affect its overall preference for the two levers, the obtained values of dB' might vary linearly with that preference (at least within the approximate range of preference of 0.45 to 0.55).

DISCUSSION

The present experiment failed to demonstrate that rats are sensitive to the rapidly-adjusting delay to reinforcement used in the task of Mazur (1987). The simulations suggested that even in the absence of such sensitivity, manipulations that affect subjects' overall preference for the two alternatives may have systematic effects on dB', the primary behavioural measure in this task. Furthermore, individual subjects did not, in general, exhibit stable patterns of choice (Figure 57, p. 153), a reason to question whether the task would be suitable for studying the effects of acutely-administered drugs on preference for delayed reinforcement. The simulations also indicated that the stability criteria previously applied to this task do not provide a guarantee that subjects are choosing other than at chance.

Interpretation of cross-correlational analysis

A number of analytical techniques were applied to this task for the first time. Analysis of the computer simulations demonstrated that correlating subjects' choices with the adjusting delay (dB) operative at the moment of choice successfully detects 'perfect' sensitivity to the adjusting delay (Figure 61), but fails to detect more complex forms of delay sensitivity such as sensitivity to a running average. The cross-correlational technique was more powerful, and successfully detected all the causal relationships embedded in the task itself (influences of choice on delay) and in the simulated decision rules (influences of delay upon choice, where applicable). When group data were considered, the cross-correlational analysis did not falsely detect causal relationships that were not present. This suggests that this technique, although complex, might be a useful way to understand the causal relationships operating in this schedule.



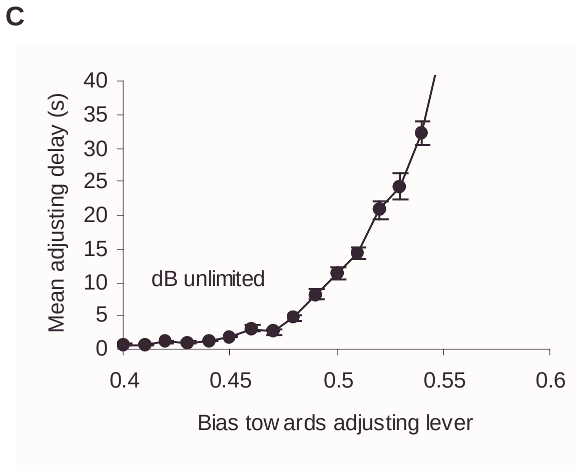
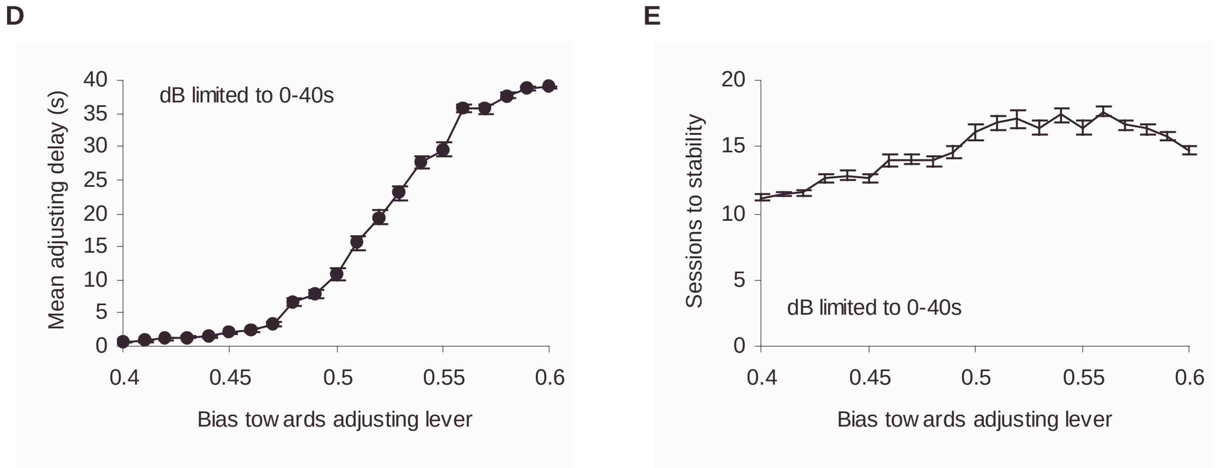


Figure 65. Effect of bias on dB' and on the number of sessions to meet the stability criteria of Mazur (1988) in simulated subjects whose behaviour was independent of the programmed delay. Every data point represents the results of 100 different simulated subjects. Panels A, B and C show the results of simulations in which no limits were placed on dB. (Panel C shows the data in Panel A, replotted with a different *y*-axis.) Panels D and E show simulations in which dB was limited to the range 0–40 s.



The technique failed to detect any consistent effect of the adjusting delay on the choices of the rat subjects.

It is not clear whether additional useful information can be gleaned from the prewhitening procedure that was applied to the data before it satisfied the assumptions of cross-correlation. For example, significant autocorrelation was only detected in the choice pattern of simulated decision rules that exhibited sensitivity to dB. (Autocorrelation was always observed in the sequence of values of dB, a consequence of the rules of the task.) Autocorrelation was, on the other hand, detected in the choice patterns of the rats, even though no influence of the dB upon choice was detectable by cross-correlation. This suggests that the rats exhibited a degree of cyclic behaviour that was unrelated to the adjusting delay. The suggestion of cyclicity is borne out by inspection of individual records in Figure 57 (p. 152); the surprising finding is that this cyclicity is apparently not a direct consequence of the adjusting delay.

Stability does not imply sensitivity to the adjusting delay

An important point that emerges from these simulations is that apparent stability cannot be taken as evidence of subjects' titrating their preference between the two alternatives. For example, some of the data series shown in Figure 60 could be taken as stable by visual inspection. Even when investigators use formalized stability criteria, delay sensitivity is not implied. As discussed on p. 158, Mazur (e.g. 1987; 1988) has used quite strict criteria to determine when a subject has reached stable performance. In addition, Mazur reduced the likelihood of finding spurious stability by taking the final value of dB from one session as the starting value for the next, rather than applying the smoothing technique used by Wogar *et al.* (1992; 1993b), in which the mean value of dB for the last half of one session is taken as the new starting point. Nevertheless, randomly-deciding simulated subjects achieved Mazur's criteria within times comparable to real pigeons (p. 163).

Possible reasons for the present failure to observe sensitivity to dB

The first explanation that must be considered is that the rats *were* sensitive to dB, but in a way that was not detected by the present analyses; perhaps the sensitivity was fleeting, or its nature changed across the course of the experiment and was masked by analysing the entire sequence of choices made by each rat. Other than the occasional cross-correlational peaks that reached significance (Figure 59A, p. 155), which were also apparent in a delay-independent simulation (Figure 63, p. 163), no evidence was found for delay sensitivity in the rats. Still another possibility is that the rats did not generalize from the forced exemplar presentations to the choice trials, and thus their preference for the adjusting alternative depended upon how often they had sampled it recently, as well as upon dB, in a highly complex feedback manner.

A more obvious explanation is that the rats were not sensitive to dB at all. Of course, the present results may not be representative of performance on this schedule generally; the lack of sensitivity may have been a consequence of procedural differences between the present experiment and previous studies. These differences may be enumerated as follows:

1. The reinforcers used were one and two 45-mg sucrose pellets. As Mazur (1987) used 2 s or 6 s of access to grain as the reinforcer for pigeons, a larger relative magnitude, it may be argued that the rats in the present study failed to discriminate between the large and small reinforcers; different results might have been obtained if the delayed reinforcer had been larger. However, at least two rat-based studies of the adjusting-delay schedule have used one and two 45-mg food pellets as the reinforcers, with 'molar' behavioural results that indicated that the subjects discriminated between them (Wogar *et al.*, 1992; 1993b).
2. In the present experiment, the adjusting delay dB was varied by 30% at a time (or 20% for the last part of the experiment). In the original studies of pigeons (Mazur, 1987; 1988), dB was altered arithmetically, typically by ±1 s; changing dB by 30% may have resulted in large swings in preference. However, proportional alterations of 30% have previously been used successfully (Wogar *et al.*, 1992; 1993b). Furthermore, Mazur (1988) has shown that increasing the step size has relatively little effect on the stable value of dB', though larger steps produce greater variability (as might be expected) and dB' sometimes increases with large step sizes.
3. The adjusting delay was not allowed to go below 2 s. In Mazur's early experiments, the floor on dB was zero; obviously, a zero floor is not possible with a proportional alteration, but it is true that

studies using proportional alterations (Wogar *et al.*, 1992; 1993b) have not placed a floor value on dB. It is possible, therefore, that the titration procedure failed because of this. If temporal discounting were steep enough that the fixed alternative (one pellet delivered after 0 s) was preferred to two pellets after 2 s, the indifference point would not be achievable, and subjects would simply keep dB at its floor value. It is possible that subjects C1, C2, and C5 attempted to do so (Figure 57, p. 153), though the CCF analysis did not demonstrate that they did, and the other five subjects certainly did not. In general, rats appear to be better able to wait for delayed reward than pigeons (see Mazur, 2000), making this interpretation less likely.

4. Similarly, a ceiling was placed on dB; initially, this was 20 s (following Wogar *et al.*, 1993b), though it was found necessary to increase this in the course of the experiment. The pigeon studies mentioned above did not place a ceiling on dB; however, as Figure 57 (p. 153) shows, once the ceiling was raised to 45 s, no rat preferred the adjusting alternative exclusively.
5. Trials were presented at constant intervals, in order to ensure that subjects could not do better by choosing 'small and often' instead of 'large but infrequent' rewards. This inevitably enforces a ceiling on dB. In several studies using this schedule (1987; Mazur, 1988), the time *between* trials was fixed; thus, the possibility existed for subjects to do well by choosing the smaller reinforcer and so being able to gain reward more often. Despite this procedural difference, systematic variations in the ITI do not appear to affect dB' in pigeons (Mazur, 1988).

Manipulations of dA and other parameters were not conducted in the present experiment, as a primary purpose of the experiment was to establish rapidly whether the task would be suitable for pharmacological and lesion studies. The present results are therefore not conclusive, as it has not been shown that the molar results of the present experiment are comparable to previous work (e.g. 1987; Mazur, 1988). In particular, it has not been demonstrated that dB' responds to long-term changes in dA in the same subjects that are insensitive to dB. However, this possibility will be explored briefly.

Effects of manipulations that alter subjects' preferences in a delay-independent manner

Effects of extrinsic manipulations

The present simulations show that evidence of an alteration in dB' as a result of a behavioural or neural manipulation is not proof of delay-dependent decision-making. Comparison of the group means from the Random and Biased-60 simulations (Figure 60, p. 160; analysis, p. 164) demonstrated that differences in relative preference for the two alternatives can lead to differences in dB', even though the decision rules generating these data took no account of dB. (This analysis illustrated a between-group difference, but the principle applies equally to a within-subjects manipulation.) Therefore, caution should be exercised when interpreting individual or group differences in dB' as an effect of a manipulation on delay sensitivity.

Effects of dA on dB'

It is clear that pigeons performing on an adjusting-delay schedule are sensitive to variations in the delay to reinforcement of the unadjusting alternative (dA) (e.g. Mazur, 1988; 1997); typical results are reproduced in Figure 66. On the basis of the present data, it is tentatively suggested that subjects performing this task are unable to track changes in dB. According to this hypothesis, they are unable to choose on the basis of the rapidly changing delay dB, and so come to assign a certain 'overall value' to the adjusting alternative. The perceived value of the unadjusting alternative, however, is constant over long periods of time, and when it changes suddenly, the 'value' assigned to the unadjusting lever changes accordingly.

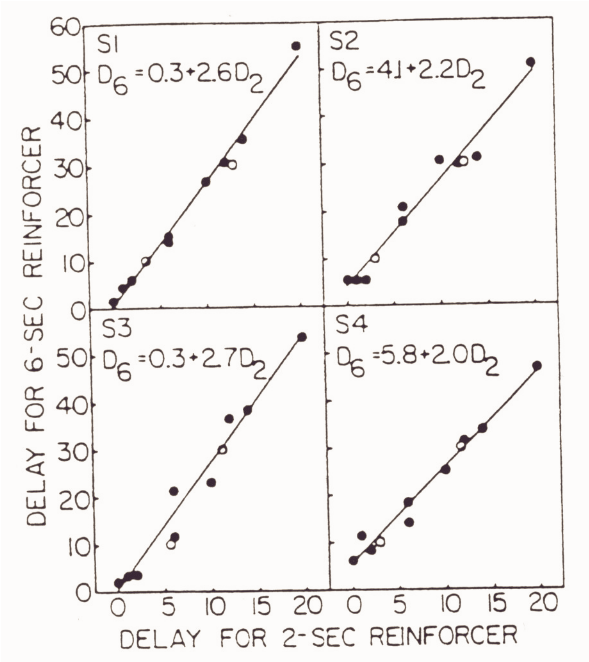


Figure 66. Alterations in dA ('delay for 2-sec reinforcer') affect dB' ('delay for 6-sec reinforcer') in an orderly fashion in four pigeons. From Mazur (1987).

On any given choice trial, subjects ignore the current value of dB but instead compare the value of the unadjusting alternative with the 'overall' value of the adjusting alternative, giving rise to a dB-independent preference. The results of the simulations depicted in Figure 65 (p. 165) show that this relative preference may be translated into a quasi-stable value of dB', and that preferences within a certain range (approximately 45–55% preference for either alternative) are related near-linearly to the value of dB'. The rats in the present study made 52.6% of free-choice responses on the adjusted lever on average (range 47.0–56.5%, SD 4.1%), clearly in the range in which a manipulation affecting relative preference could alter dB'. In summary, this hypothesis states that subjects are sensitive to dA but not directly to dB. While this may not be an appealing idea, it seems possible.

Indeed, it has been observed that bias for the adjusting alternative, measured as the ratio of dB' to dA when the two reinforcers are equal ('bias for or against the adjusting procedure itself'; Mazur, 1984, p. 429), increases as a function of dA (Mazur, 1984, p. 431; though see Mazur, 1987, p. 63; Mazur, 1988, p. 46).

In principle, similar arguments regarding manipulation effects and bias apply to adjusting-magnitude tasks. The adjusting-magnitude task (Richards *et al.*, 1997b) is learned faster than the adjusting-delay task (see Ho *et al.*, 1999, p. 369), suggesting that rats may learn the contingencies more readily with varying reinforcer magnitudes than with varying delays to reinforcement (an interpretation compatible with studies of instrumental learning with delayed reinforcement, e.g. Lattal & Gleeson, 1990; Dickinson *et al.*, 1992). However, the adjusting-magnitude task also involves a titration method in which subjects' preference affects a variable that is assumed to affect subjects' preference in turn.

To emphasize a point, the present simulations do not prove that subjects in previous studies were insensitive to dB, but demonstrate that many of the observed molar features of performance can be obtained in the absence of such sensitivity.

Comparison to free-operant schedules of reinforcement

In contrast to the present experiment, rats are known to be able to track reinforcement rate extremely rapidly in some circumstances, and probably do so by timing interreinforcement intervals (IRIs). This was demonstrated by Mark & Gallistel (1994), who used two concurrent variable interval (VI) schedules of lateral hypothalamic stimulation, ranging from VI 4 s to VI 256 s. They showed that the rats' response allocation tracked not only changes in the programmed ratio of reward between the two levers, but also the unprogrammed random fluctuations in the VI schedule, to an extent that their behaviour was governed by a very few of the most recent IRIs. This result implies that rats do *not* maintain and use a decaying 'running average' of the reward history, at least in that task (see Mark & Gallistel, 1994, pp. 90–91); Mark & Gallistel argue persuasively that their rats tracked the relative ratio of reward rate on the two levers by timing the interval to detect a fixed number of rewards (this number being from one to three).

It is an interesting question as to why rats are apparently capable of timing intervals on a seconds-to-minutes timescale and updating choice behaviour based on these intervals in concurrent VI schedules, but are apparently incapable of this in the discrete trials adjusting-delay procedure. It must be acknowledged that the two procedures are very different. Discrimination of changes in relative reinforcement rate may be easier than discrimination of changes in reinforcement delay in a discrete-trial procedure. One possibility, discussed by Mark & Gallistel (1994, p. 94) is that regular, dramatic changes in reward encourage extreme sensitivity to these changes, while relative stability with slow changes in reinforcement parameters (as in the present task) discourages local sensitivity to the reinforcement contingencies. Whether this reflects the operation of two psychological processes is unclear, but relative invariance of response–reinforcement contingencies has been suggested to be the key factor engendering habitual responding (Dickinson, 1985) (as discussed in Chapter 1, p. 25); discrete-trial schedules constrain behavioural variability much more than free-operant schedules. One highly speculative interpretation is that the task of Mark & Gallistel (1994) tests goal-directed action while choice in Mazur's (1987) procedure is more heavily influenced by the relative strength of two differentially reinforced stimulus–response habits.

SUMMARY

The adjusting-delay task has produced consistent results on the molar scale and lends itself well to using dB' values as a measure of relative preference of different 'fixed alternative' conditions (work reviewed by Mazur, 1997). However, caution must be exercised when interpreting effects on dB' as changes in sensitivity to dB. In the present study, rats did not update their behaviour rapidly to reflect changes in dB, and no clear evidence for any form of sensitivity to dB was found. These results suggest that rats' behaviour on this task would not be characterized well as 'informed choice'; the psychological mechanisms underlying choice in this task are not clear at present. Artificial decision rules that take no account of dB were found to be able to replicate a number of observed features of performance on the task, including the satisfaction of stability criteria and the generation of within- or between-subject differences in dB'. Finally, rats' preference did not exhibit clear stability or consistency even after prolonged training. The task therefore appears unsuitable for acute pharmacological studies, for which it would be preferable to be able to perturb and re-stabilize performance within one or a few sessions. In Chapter 6, a task of the 'systematic' kind will therefore be turned to, in which the subject has no influence on the delay to reinforcement.