*STOCHASTIC CHOICE MODELS: A COMPARISON BETWEEN
BUSH-MOSTELLER AND A SOURCE-INDEPENDENT
REWARD-FOLLOWING MODEL*

J. E. R. STADDON AND J. M. HORNER

DUKE UNIVERSITY AND COLORADO COLLEGE

Horner and Staddon (1987) argued that a class of reward-following processes defined by a property they termed *ratio invariance* is a better model for the probabilistic choice performance of pigeons than competing molecular accounts such as momentary maximizing, melioration, and the Bush-Mosteller model. The critical data were provided by choice distributions—distributions of a variable S , the proportion of Right choices, defined on a moving window typically 32 choices long—obtained under a *frequency-dependent* schedule. The schedule prescribed equal payoff probabilities, $p(S)$, for both choices. $p(S)$ was a maximum when $S = 0.5$ and declined linearly for S values above and below 0.5. Pigeons showed generally bimodal choice distributions with the modes at equal $p(S)$ values. These data do not follow easily from melioration or momentary maximizing and are inconsistent with molar maximizing, but they may be consistent with Bush-Mosteller. We present here the results of computer simulations showing that the ratio-invariance model studied yields, as expected, choice modes at equal $p(S)$ values, but that Bush-Mosteller, although capable of generating bimodal choice distributions, does not have choice modes at equal $p(S)$ values.

Key words: dynamic model, quasi-dynamic model, probabilistic choice, linear operator, simulation, ratio invariance, pigeons

The law of effect is deceptively simple. In its modern form it asserts just that reinforcement (reward) increases the probability of the rewarded response, that is, the response with which it is contiguous. The obvious question, which has energized the field of animal learning for many years, is *by how much* does a reward increment response probability? The most popular answer has been provided by the linear operator model first proposed many years ago by Bush and Mosteller (1955) and extended by a number of others, most notably Rescorla and Wagner (1972; see also Widrow & Hoff, 1960). Largely for reasons of mathematical convenience, Bush-Mosteller assumes that each reinforcement increments response probability by an amount that is a constant fraction of the difference between the current probability and the maximum probability (i.e., 1.00). Nonreinforcement acts in

the same way, reducing response probability by a fraction of the difference between current response probability and the minimum (i.e., 0). In symbols, for a two-choice situation in which s refers to the probability of a Right response:

For a Right response, following:

$$\begin{array}{l} \text{Reinforcement (R)} \\ \delta s = A(1 - s) \end{array} \quad (1)$$

$$\begin{array}{l} \text{Nonreinforcement (N)} \\ \delta s = -Bs, \quad 0 < B < A < 1, \end{array} \quad (2)$$

where s is response probability [Right/(Left + Right)], δs is the change in s after R or N, and A and B are constants. A symmetrical pair of equations describes the changes following reinforcement and nonreinforcement for a Left response. All four δs functions are shown in Table 1. A and B are assumed to depend on such things as deprivation level, reward size, and reinforcement history.

This model, although probabilistic with respect to s , is completely deterministic with respect to changes in s . The linearity of the model allows us to aggregate across individuals so as to derive deterministic predictions for such things as average acquisition and extinction curves (cf. Sternberg, 1963). Moreover, if the

This research was supported by grants to Duke University from the National Science Foundation and The Pew Memorial Trust. J. Staddon thanks the Alexander von Humboldt-Stiftung for support and Juan Delius and the Ruhr-Universität, Bochum, Federal Republic of Germany, for hospitality during the preparation of this paper. Correspondence and reprint requests may be sent to J. E. R. Staddon, Department of Psychology, Duke University, Durham, North Carolina 27706.

Table 1

Bush–Mosteller Model for the two-arm bandit situation. In each quarter-panel, the outcome probabilities are at the top (reward probabilities are p and q on Right and Left; choice probability $s_n = R_n / (R_n + L_n)$), and the changes in choice probability, δs (i.e., $s_{n+1} - s_n$), are at the bottom in brackets.

RESPONSE ON:	Left	(1- s_n) q [- As_n]	(1- s_n)(1- q) [$B(1-s_n)$]
		Reward	Nonreward
	Right	s_np [$A(1-s_n)$]	$s_n(1-p)$ [- $B(s_n)$]
		OUTCOME	

probability of reinforcement for each alternative is constant (concurrent-ratio schedule, or two-armed bandit) we can easily derive predictions for the variance of the distribution of choices and the matrix of first-order transition probabilities. Predictions of this sort have been tested in a number of situations over the years. The results have been mixed: Sometimes acquisition and extinction curves show the exponential form predicted by Bush–Mosteller, but often they do not; sequential statistics similarly often fail to follow the predicted pattern (e.g., Atkinson, Bower, & Crothers, 1965; Sternberg, 1963). The details of choice performance, and transitions from one asymptotic condition to another (e.g., extinction and acquisition curves), seem to be more varied than the possibilities permitted by Bush–Mosteller.

The greatest successes of the Bush–Mosteller model, both in its original version and in the extended form represented by the Rescorla–Wagner (1972) model, have been its ability to account for asymptotic (i.e., equilibrium) performance. But in application to the molecular details of behavior and to transition states, the model in a sense predicts too much. The data are often highly variable and differ both within and between animals, yet Bush–Mosteller is constrained always to predict uniform results.

These considerations led us to explore a number of alternatives to Bush–Mosteller. Each alternative was defined by a particular rule for δs . For example, an early class of models

assumed that the effects of R and N were simply constant, that is,

For a Right response, following:

Reinforcement (R)

$$\delta s = a \quad (3)$$

Nonreinforcement (N)

$$\delta s = -b, \quad 0 < b < a < 1, \quad (4)$$

where a and b are constants. (Of course, s had to be limited in some way to the interval $0 \leq s \leq 1$). The four possible values for δs implied by this model are shown in Table 2.

This simple constant-difference (CD) model worked quite well to explain some otherwise puzzling variations in the equilibrium properties of performance in the two-armed bandit situation and, subsequently, on frequency-dependent schedules (Horner, 1986; Staddon, 1988). However, it suffers from the disadvantage that it predicts a fixed and rather improbable straight-line form for the highly variable acquisition and extinction functions. Like Bush–Mosteller, it predicts too much.

Comparing Table 2 with Table 1 shows that Bush–Mosteller and the CD model differ in a fundamental way: For Bush–Mosteller, the magnitude of the effect of reward or nonreward depends on which response is rewarded or not rewarded, whereas for the CD model the magnitudes of the reward and nonreward effects are the same for both responses. For the CD model, reward always increases the probability of the rewarded response by the quantity a , nonreward decreases it by b . We called this property *source independence*. We have argued elsewhere as to the relative plausibility of this assumption compared to the Bush–Mosteller assumption that the effects of reward and nonreward depend on the baseline response probability (Horner & Staddon, 1987; Staddon, 1988). Here, we focus on the fact that it is possible to generalize the idea of source independence as shown in Table 3.

The generalized source-independent model is little more than a quantitative restatement of the verbal law of effect: Reward increases response probability, nonreward decreases it. The amounts of change are unspecified beyond the assumption that they depend in some way on s . All that is added is the idea that in a symmetrical two-choice situation, the expected values of the changes will be the same for each response. This addition is not sufficient to al-

Table 2

Expected changes δs ($=s_{n+1} - s_n$) in choice probability associated with the four possible outcomes in a two-choice experiment, according to the constant-difference model.

RESPONSE ON:	Left	-a	b
	Right	a	-b
		Reward	Nonreward
OUTCOME			

low us to derive predictions, however, even in the simplest case (the two-armed bandit). Prediction becomes possible if the relation between $a(s)$ and $b(s)$ is constrained in certain ways. For example, if the ratio $a(s)/b(s)$ is assumed to be constant (we term this *ratio invariance*: RI) it becomes straightforward to derive predictions about equilibria, even though the fact that $a(s)$ and $b(s)$ are unspecified makes it impossible to predict anything about transient properties of the reward-following process.

Notice that the CD model (i.e., R and N produce fixed changes in choice probability) is a special case of the RI model (if a and b are constant, so is their ratio).

Horner and Staddon (1987) derived equilibrium predictions from ratio invariance for three situations: two-armed bandit (p, q constant), a symmetrical frequency-dependent (SFD) schedule ($p = q$, and both dependent on current choice proportion, S), and an asymmetrical frequency-dependent schedule ($p \neq q$, and both dependent on S). Their first ex-

Table 3

Generalized version of source-independent reward-following. The changes in s caused by reward and nonreward can now be any function of response probability, so long as they are the same for both responses (i.e., $a \rightarrow a(s)$ and $b \rightarrow b(s)$).

RESPONSE ON:	Left	-a(s)	b(s)
	Right	a(s)	-b(s)
		Reward	Nonreward
OUTCOME

Fig. 1. Symmetrical frequency-dependent schedule. The bilinear function shows the reward probability, $p(S)$, for both responses as a function of choice proportion, S [$= R/(R + L)$], measured in a moving window of M responses. The two dotted lines define the range of w values for which there are partial-preference modes on both sides of indifference. The dashed horizontal line shows the choice modes (at S_1 and S_2) predicted for a single w value. See text for details.

periment focused on the SFD schedule, defined as follows. Current choice proportion, S , was measured in a continuously updated (moving) “window” of M responses.¹ For example, if M is 4, the current value of S for the following choice sequence, LLLRLRL, is 0.5; if the next response is L, the value of S changes to 0.25. In our experiments, M varied from 4 to 96, 32 being the value we used most frequently. The value of reward probability was always the same for both choices: $p(S) = q(S)$, but $p(S)$ depended on S according to the bilinear function shown in Figure 1, that is, reward probability was maximal for both responses when $S = 0.5$ (i.e., indifference), and then decreased, at different rates, at S values greater or less than 0.5.

Frequency-dependent schedules play a critical theoretical role, because they allow us to discriminate between four theories—maximizing (momentary and overall), matching (melioration), and ratio invariance—that under many other conditions give rise to very similar predictions. Because payoff probability on the SFD schedule is always the same for both alternatives (so that matching is forced),

¹ Note that the *schedule window* (the value of which we term M) need not be the same as the *data-analysis window* (m), although the two have always been the same in our published experiments. It is quite possible to record choice performance in a window of size 16, say, but to base reward probability on a window of size 64.

neither momentary maximizing nor matching makes any point prediction. The maximum $p(S)$ at $S = 0.5$ means that molar maximizing must predict indifference. But ratio invariance predicts a bimodal choice pattern, which was confirmed by our data.

There is a fifth theoretical alternative, however, the Bush–Mosteller model, that was not ruled out by our previous analysis. Bush–Mosteller, like RI, also predicts a bimodal choice pattern. The present note is intended to clear up this ambiguity by showing (a) that the bimodal choice distribution predicted by Bush–Mosteller differs in critical ways from that predicted by RI, and (b) that the data conform more closely to the RI prediction.

PREDICTIONS FOR THE SYMMETRICAL FREQUENCY-DEPENDENT SCHEDULE

Ratio Invariance

We derived equilibrium predictions of the RI model for the SFD schedule, $p(S) = q(S)$, in the following way (see Horner & Staddon, 1987, for the full account). From the basic equations (Table 3), we can derive the expected value for δs , delta (s), as a function of s and S which, neglecting a constant multiplier, is

$$\delta s = 2s[p(S) - w] + w - p(S), \quad (5)$$

where s is choice probability (a theoretical value), S is choice proportion (an empirical value, measured across a moving window M choices long), and w is the ratio of nonreward to reward effects ($w = b(s)/[a(s) + b(s)]$). To derive equilibrium predictions, we then made the further assumption that S , the measured choice proportion, is an unbiased measure of s , choice probability, which allows S to be substituted for s on the right-hand side and permits Equation 5 to be solved once $p(S)$ and $q(S)$ are specified.

The critical property of the equilibrium points is obvious from Equation 5, which is zero whenever $p(S) = w$. For any linear SFD schedule $p(S) = Ks$, the equilibrium at $S = w/K$ is stable (and the other equilibrium at $S = 1/2$ unstable) if $w < |K/2|$. With a bilinear SFD schedule satisfying this condition (such as the one in Figure 1), there are therefore two possible stable modes, which must be at

equal $p(S)$ values.² Thus, if w is at the value indicated by the dashed horizontal line in Figure 1, stable equilibria are possible at the two S values shown, S_1 and S_2 . A critical test for RI, therefore, is to see if, when the data show two partial-preference modes (i.e., two modes both at S values other than 0 or 1), these modes are at equal $p(S)$ values. Horner and Staddon (1987) confirmed this prediction and concluded that RI provides a better account of choice in these purely probabilistic situations than competing theories such as melioration, momentary maximizing, and the Bush–Mosteller model.

Bush–Mosteller

But it turns out that matters are not quite this simple. Although models such as melioration and momentary maximizing do indeed seem to be ruled out by our data, the Bush–Mosteller model is not. Donald Laming (personal communication) pointed out that at high reward probabilities (or high relative values for parameters A and B in Equations 1 and 2) Bush–Mosteller implies that successive choices become highly correlated with one another. Such an autocorrelation, filtered through our moving-window method of data analysis, can lead to the bimodal patterns we found. This is obvious to intuition. If A is high relative to B, for example, then a single reward will shift s close to one extreme or another, which will tend to cause a rewarded choice of Left or Right to be repeated. Runs of L or R choices, when averaged through a window, will tend to show two modes whose locations will depend on the size of the data-analysis window, m . The smaller the value of m , the closer to the extremes, 0 or 1, will be the modal values of the preference distributions (in the limit, when $m = 1$, only S values of 0 or 1 are possible).

Although the Bush–Mosteller predictions for the SFD procedures cannot easily be de-

² Note that there is no contradiction between our assertion that a mode is stable and our data showing more than one mode. *Stability* is defined with reference to perturbations. In this case perturbations are provided by the stochastic $p(S)$ schedule. Each mode is stable under small perturbations, but the system can pass from one mode to another if $p(S)$ varies sufficiently. Figure 3 in Horner and Staddon (1987) shows the argument graphically and Figure 4 shows an empirical example of such a transition.

rived analytically, it is possible to derive predictions via computer simulation. That bimodal choice distributions can be produced is obvious. What is not obvious is whether or not the peaks of these distributions will be at equal $p(S)$ values, as in the data and as predicted by RI. Answering this question, for Bush–Mosteller and for a dynamic version of the RI model, was the major objective of our simulations.

SIMULATIONS

Ratio Invariance

In order to produce choice-by-choice predictions from ratio invariance it is necessary to decide on a specific form for the two δs functions, $a(s)$ and $b(s)$. We looked at three different possibilities ($b(s)/[a(s) + b(s)] = w = \text{constant}$, and $a(s) > b(s) > 0$, in all cases):

1. $a(s) = a$, $b(s) = b$, $0 < b < a \ll 1$, that is, reward (and nonreward) change s by a small, constant amount. The boundary ($0 \leq s \leq 1$) is maintained by setting s values greater than 1 equal to 1, and values less than 0 equal to 0.

2. As above, except that R and N affect directly a dummy variable s^* , which is allowed to take values in the interval $-a \leq s^* \leq 1 + a$. $s = s^*$ for values in the closed interval 0–1, with larger and smaller values set equal to the extremes (i.e., $s = 1$ for $s^* > 1$; $s = 0$ for $s^* < 0$)—and similarly for b . This allows for large values of a and b , but still ensures that s is within the 0–1 interval.

3. $a(s) = s(1 - s)a$, $b(s) = s(1 - s)b$; so long as a and b are not too large, this transformation forces s to be within the 0–1 interval.

Model 1 turned out to be impractical because, when a and b are sufficiently small to minimize “end effects,” the value of S changes only very slowly. Model 3 has some complications we do not yet fully understand. We therefore here present data only from Model 2. In all cases M , the schedule-window size, and m , the data-analysis window, were set equal to 80. We looked at five values for a and b , yielding 25 different simulations and nine different values of w (Table 4) for each. For each of these a, b pairs we obtained at least 10^6 choice responses and generated choice-frequency distributions like the one shown in Figure 2. For Model 2, 19 of these distributions

Fig. 2. Choice frequency distribution from the ratio-invariance simulation (Model 2). The bilinear function is the frequency-dependent schedule; it shows the reward probability for both responses at various values of S (choice proportion) measured in a moving window 80 responses long. The choice-frequency distribution (in arbitrary units) summarizes 10^6 choices generated according to Model 2, with $a = 0.125$ and $b = 0.0125$. Note that the modal choice proportions correspond to approximately equal values of $p(S)$, as theory requires.

turned out to be bimodal. In Figure 3 the filled triangles show the values of $p(S)$ corresponding to the two modes, plotted against each other. As theory implies, so long as both modes represent partial preferences, they are at approximately equal $p(S)$ values. The horizontal line of points to the left at $p(S) = 0.04$ represents those cases in which one of the modes was at exclusive choice of the Right response, which corresponds to a $p(S) = 0.04$ (cf. Figure 2). When neither mode is at exclusive choice, the points all lie reasonably close to the unit-slope line, as theory requires.

For purposes of comparison, we also show data from Horner and Staddon (1987), plotted as open circles in Figure 3.³

Figure 4 shows the relation between the S values of the two modes from the Model 2 simulation. The graph plots the absolute difference between the modal value and 0.5; thus if both modes were always disposed symmetrically about indifference, the resulting set of

³ Because of programming limitations, the left limb of the bilinear reward function used in the Horner and Staddon (1987) experiment did not pass through the point (0,0). In the simulations, the reward function was $p(S) = 0.16S$; in the experiment the reward function was $p'(S) = 0.16S + 0.005$, that is, $p'(S) = p(S) + 0.005$. To allow visual comparison of our data with the results of the simulations we have therefore subtracted 0.005 from each obtained $p(S)$ value.

Fig. 3. Modal $p(S)$ values for bimodal choice distributions from the simulations of ratio invariance, Model 2 (solid triangles). The $p(S)$ values of the two modes are plotted against one another. Solid line is the locus of equality. Open circles are transformed data (see text) from Horner and Staddon (1987).

Fig. 4. Modal S values for bimodal choice distributions from the simulations of ratio invariance, Model 2 (same data set as Figure 3). The absolute differences between 0.5 and the S values of the two modes are plotted against one another (solid triangles). Solid line indicates equality; dashed line through the origin has a slope of 2. Open circles are transformed Horner and Staddon (1987) data.

points would fall on a straight line of unit slope through the origin. In fact the points fall on a line of slope equal to two, until the right-hand mode reaches the extreme value of unity. Figures 2 and 3 are not independent, of course, but linked by the bilinear schedule function (Figure 1). As before, the open circles show the Horner and Staddon (1987) data.

Bush–Mosteller

We carried out 25 Bush–Mosteller simulations, with the parameter values shown in Table 5. Of these 25, 13 produced bimodal choice distributions, with no modes at exclusive choice. Figure 5 shows a typical example. The open triangles in Figure 6 show how the $p(S)$ values of these 13 mode-pairs were related. They follow a curvilinear path well above

the line of equality. Figure 7 shows the corresponding relation between the modal S values. These points lie close to the line of equality. The filled circles show the Horner and Staddon (1987) data in both cases.

Effect of M Value

In all our simulations and data analyses, the data-analysis window, m , was the same as the schedule window, M . There is an element of arbitrariness in this convention. How are our conclusions—particularly the conclusion that the peak $p(S)$ values for bimodal RI distributions are equal—affected by changing M from the value of 80 used in the simulations? We attacked this problem by repeating one RI simulation ($a = 0.128$, $b = 0.008$) using an M

Table 4
Parameter values used for the ratio-invariance (CD) simulations. Cell entries are $w = b/(a + b)$.

Parameter a	Parameter b				
	0.01250	0.01000	0.00800	0.00640	0.00512
0.20000	0.0588	0.0476	0.0385	0.0310	0.0250
0.16000	0.0725	0.0588	0.0476	0.0385	0.0310
0.12800	0.0890	0.0725	0.0588	0.0476	0.0385
0.10240	0.1088	0.0890	0.0725	0.0588	0.0476
0.08192	0.1324	0.1088	0.0890	0.0725	0.0588

Fig. 5. A typical bimodal choice distribution generated by Bush-Mosteller in conjunction with the bilinear symmetrical frequency-dependent schedule. Other information as in Figure 1.

value of 10, but leaving the data-analysis window, m , at 80. The results are summarized in Table 6, which shows the modal S and $p(S)$ values under the two conditions. Although M varies from 10 to 80 and the choice modes shift towards indifference at the lower M value, as expected, the $p(S)$ values for the two choice modes remain approximately equal.

CONCLUSION

Simulation is not as satisfactory as analysis as a way of understanding the properties of a process because the inferences we can make

Fig. 7. Modal S values from bimodal choice distributions from the simulations of Bush-Mosteller (open triangles). Filled circles are data from Horner and Staddon (1987). Other information as in Figure 4.

are always limited to the parameter values we have tried. Analysis permits generalizations true for *all* parameter values within a given range. Consequently, the conclusions we can draw from the present set of simulations are more tentative than we would like. Nevertheless, the following conclusions seem warranted: (a) Bush-Mosteller does indeed yield bimodal choice distributions under the SFD schedule under many conditions; but (b) the modes do not generally correspond to equal $p(S)$ values. (c) The ratio-invariant CD model more easily yields bimodal distributions, including distributions in which one mode is at exclusive choice. When neither mode is at exclusive choice, both modes correspond to similar $p(S)$ values. In all these respects the CD model is a better fit to the data than is Bush-Mosteller. Finally, (d) the relation between the two modal $p(S)$ values does not depend critically on the two averaging windows: We found a similar relation when the data-analysis window and the schedule window were the same as when they were different—although the absolute location of the modes changed. Hence, our earlier theoretical assumption that S (measured choice proportion) is an unbiased estimate of s (the theoretical value of choice probability) seems justified.

To researchers not intimately involved in the minutiae of stochastic processes, the ques-

Fig. 6. Modal $p(S)$ values for bimodal choice distributions from the Bush-Mosteller simulations (open triangles). Filled circles are data from Horner and Staddon (1987). Other information as in Figure 3.

Table 5

Parameter values used for the Bush–Mosteller simulations. Cell entries are $B/(A + B)$.

Parameter A	Parameter B				
	0.00400	0.00200	0.00100	0.00050	0.00025
0.400	0.00990	0.00498	0.00249	0.00125	0.00062
0.200	0.01961	0.00990	0.00498	0.00249	0.00125
0.100	0.03846	0.01961	0.00990	0.00498	0.00249
0.050	0.07407	0.03846	0.01961	0.00990	0.00498
0.025	0.13793	0.07407	0.03846	0.01961	0.00990

tions we have addressed here may seem of little more than technical interest. We believe they have wider implications, however. Learning psychologists have over the years comfortably assumed that deciding between various dynamic models of learning would not be difficult; the difficulty, it was believed, was in coming up with promising possibilities. This note illustrates the point that deciding among the large number of potential reward-following processes animals can use will not be easy, even if we restrict ourselves entirely to probabilistic (i.e., nontemporal) situations. The properties even of processes as simple as the ones we have considered here are not always obvious. Nor is the relation between quasi-dynamic (i.e., equilibrium) and truly dynamic analyses always clear. For example, equilibrium analysis shows the RI model always to have partial-preference modes at equal $p(S)$ values. Our simulations with $a(s)$ and $b(s)$ both constant confirm this. Yet when we multiply both by the quantity $s(1 - s)$ (as in Model 3 above) which does not change the equilibrium predictions, the much longer “stay” times at extreme values of s , when $a(s)$ and $b(s)$ are both necessarily very small, mean that measured modes are actually almost always at exclusive choice. We conjecture that the equilibrium

predictions are only correct, therefore, if the expected values of $a(s)$ and $b(s)$ are independent of s , but rigorous proof is currently lacking. Similarly, very small values for $a(s)$ and $b(s)$ produce a “sticky” choice performance that shows only a single mode, strongly dependent on the initial value of s . Here the equilibrium predictions are certainly correct, but an impracticably large number of iterations is required for a valid simulation. In short, although the concept of reward following (i.e., the law of effect) is not in doubt, discovering the details of the process presents a number of subtle difficulties that are yet to be fully addressed by learning theorists.

REFERENCES

Atkinson, R. C., Bower, G. H., & Crothers, C. J. (1965). *An introduction to mathematical learning theory*. New York: Wiley.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. New York: Wiley.

Horner, J. M. (1986). *Probabilistic choice in pigeons*. Unpublished doctoral dissertation, Duke University, Durham, NC.

Horner, J. M., & Staddon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Processes*, **15**, 59–92.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.

Staddon, J. E. R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior*, **49**, 303–320.

Sternberg, S. (1963). Stochastic learning theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 1–120). New York: Wiley.

Widrow, G., & Hoff, M. E. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record Part 4*, 96–104.

Table 6

Effect of different schedule (M) windows. Data-analysis window, m , = 80 for both. Cell entries are S values and $p(S)$ values for the two modes.

		S_1	S_2	$p(S_1)$	$p(S_2)$
M	10	.3857	.8750	.054	.050
	80	.2875	.9375	.046	.045