HILL-CLIMBING BY PIGEONS

John M. Hinson and J. E. R. Staddon

DUKE UNIVERSITY

Pigeons were exposed to two types of concurrent operant-reinforcement schedules in order to determine what choice rules determine behavior on these schedules. In the first set of experiments, concurrent variable-interval, variable-interval schedules, key-peck responses to either of two alternative schedules produced food reinforcement after a random time interval. The frequency of food-reinforcement availability for the two schedules was varied over different ranges for different birds. In the second series of experiments, concurrent variable-ratio, variable-interval schedules, key-peck responses to one schedule produced food reinforcement after a random time interval, whereas food reinforcement occurred for an alternative schedule only after a random number of responses. Results from both experiments showed that pigeons consistently follow a behavioral strategy in which the alternative schedule chosen at any time is the one which offers the highest momentary reinforcement probability (momentary maximizing). The quality of momentary maximizing was somewhat higher and more consistent when both alternative reinforcement schedules were time-based than when one schedule was time-based and the alternative response-count based. Previous attempts to provide evidence for the existence of momentary maximizing were shown to be based upon faulty assumptions about the behavior implied by momentary maximizing and resultant inappropriate measures of behavior.

Key words: concurrent schedules, optimal behavior, momentary maximizing, strategies, variable-interval, variable-ratio, key peck, pigeons

There is some argument about whether animals always perfectly maximize average reinforcement rate on simple reinforcement schedules, but all agree that they do pretty well. Optimal behavior on concurrent ratio schedules implies exclusive choice, and pigeons respond almost entirely on the schedule with the smaller average ratio (e.g., Herrnstein & Loveland, 1975). The best strategy on concurrent variable-interval, variable-interval (concurrent VI VI) is non-exclusive choice with frequent switching between schedules, and this is what animals show. There is less agreement on how animals achieve their close-to-optimal performance (Heyman & Luce, 1979; Rachlin, Green, Kagel, & Battalio, 1976).

The least plausible possibility is that animals somehow compute overall average (molar) reinforcement rates, compare the rates obtained by different moment-by-moment (molecular) patterns of choice, and then settle on the pattern that gives the highest molar rate of payoff (cf. Herrnstein & Vaughan, 1980; Heyman & Luce, 1979). The objections to this molar comparison strategy are numerous. The differences in molar reinforcement rate associated with different strategies are often trivial; to detect them in this way would require not only an exceedingly precise assessment of average rates but also a capacious and error-free memory to allow for the relevant comparisons. The theory also assumes the animal to be averaging over the same time periods as the experimenter (i.e., just from one experimental session to the next, ignoring events in between). It seems obvious that some much more limited set of processes underlies performance on concurrent intervals and ratio schedules.

This conclusion should not surprise: No animal is omniscient, and optimal behavior—in animals, people, or intelligent machines—is always the outcome of a set of processes, unintelligent in themselves, that nevertheless suffice

to do the job. One such simple process used to approximate more sophisticated maximizing outcomes is termed *hill-climbing* (cf. Minsky, 1961), because it can be illustrated by the metaphor of a blind man looking for the top of a hill. His best strategy is simply to judge the slope in different directions from his current position, then take the direction of steepest slope (another term for hill-climbing is the *method of steepest ascent*). This process will indubitably get the man to the top of something, although it may be a hillock rather than a hill—the method finds *a* peak, not necessarily the highest peak.

Shimp (1966, 1969) some years ago proposed that animals on concurrent schedules follow a hill-climbing strategy that he termed *momentary maximizing*. He suggested that molar choice behavior can be analyzed into sequences of particular choices, each of which follows a simple rule: Pick the alternative with the highest payoff probability. Despite a number of supporting experimental papers, Shimp's attractive theory never gained wide acceptance, perhaps because it seemed much more complicated to test than it was to state. In later work, mistaken comparisons between superficially similar, but actually quite different, experimental procedures have been common, largely due to ignorance of the schedule-feedback properties. In the absence of a firm theoretical foundation, tests of the theory have confused rather than clarified (Staddon, Hinson, & Kram, 1981).

There have been two major problems in the study of momentary maximizing: (a) determining the formal criterion for momentary maximizing, and (b) determining behavioral measures that accurately reflect the presence or absence of momentary maximizing. Tests of the theory seemed to require tables of sequential-response probabilities (e.g., Heyman, 1979; Nevin, 1979; Silberberg, Hamilton, Ziriax, & Casey, 1978) that are as difficult to interpret as they are inconvenient to compile.

It turns out that the test for momentary maximizing is quite simple. For example, reinforcement probability for a response made to one of two independent constant-probability variable-interval schedules depends only upon the time since the last occurrence of that response (see Staddon et al., 1981). Each choice can be defined in terms of the times since the previous such choice and the previous alternative choice. Consequently, a point in a clock space defined by these two times can be uniquely identified as consistent, or inconsistent, with momentary maximizing (we return to this analysis in a moment).

We show in this paper that pigeons on concurrent variable-interval, variable-interval and concurrent variable-ratio, variable-interval schedules conform quite well to momentary maximizing. We begin by considering several measures of momentary maximizing on concurrent variable-interval, variable-interval schedules. We go on to examine how the quality of momentary maximizing changes over time and with changes in schedule values. Finally, we compare performance on concurrent variable-interval, variable-interval schedules with performance on concurrent variable-ratio, variable-interval schedules.

EXPERIMENT 1
CONCURRENT VARIABLE-INTERVAL, VARIABLE-INTERVAL SCHEDULES

METHOD

Subjects

Six adult male White Carneaux pigeons served; each bird was maintained at 80% of its free-feeding weight. All had previous experience with various schedules other than concurrent schedules.

Apparatus

All experiments were conducted in a standard aluminum and Plexiglas operant-conditioning chamber with internal dimensions of 37 by 31 by 33 cm. The two translucent pecking keys were 2 cm in diameter, 26 cm from the floor, and 15 cm apart. Each key was transilluminated by a 6-W light. Between the two keys and equidistant from them was a 4 by 5-cm aperture to a food magazine. Recorded key pecks were accompanied by an audible click from a small relay. The reinforcer was 3-sec access to mixed grain. During reinforcement, the response keylights were extinguished and another 6-W light above the food hopper was turned on. The entire experimental chamber was enclosed by a larger, soundproofing box; the interior was illuminated by a 10-W fluorescent lamp. White noise and a small ventilation fan installed in one wall masked extraneous sounds. The experimental contingen-

cies and data recording were carried out by a microcomputer in an adjacent room. Data on the absolute time (to one msec) and identity of each experimental event were later transferred to a PDP 11 computer for analysis.

Procedure

The experimental conditions for all animals appear in Table 1. Two sets of three birds received different training conditions. One set of birds (CO96, CO123, and CO104) received training with four conditions to look for long-term effects of exposure to a given concurrent VI = x VI = y schedule. A second set of birds (CD129, CD117, and CD148) received exposure to a wider range of more frequently shifted x and y values. No changeover delays were used in any condition. Interreinforcement intervals for the VI schedules were generated using the constant-probability progression suggested by Fleshler and Hoffman (1962). Each session lasted one hour, excluding the time taken by reinforcer delivery. Some preliminary data from one bird in these experiments appeared in Hinson and Staddon (1981).

Method of Data Analysis

Shimp's analysis of momentary maximizing for concurrent variable-interval, variable-interval schedules is the most widely cited, but it is appropriate only for a particular type of discrete-trials procedure and has been wrongly applied to free-operant concurrent VI VI schedules. Perhaps the most widespread error is to assume that a fixed sequence of choice or switches is required by momentary maximizing (we return to this point in the final Discussion).

The only determinant of reinforcement probability on a constant-probability variable-interval schedule is the time since the last response to that schedule, regardless of whether or not the response resulted in reinforcement (see Staddon et al., 1981). Considering the simplest case of a single constant-probability variable-interval schedule, reinforcement probability for a response made at time t is described by

$$P(R|t) = 1 - e^{-\lambda t}, \qquad (1)$$

where λ represents the rate of VI reinforcement and e is the base of natural logarithms. For two VI schedules, momentary maximizing dictates that with every choice to respond, the response be made to the higher-probability alternative of

$$P(R|t_1) = 1 - e^{-\lambda_1 t_1} \qquad (2)$$

$$P(R|t_2) = 1 - e^{-\lambda_2 t_2}, \qquad (3)$$

where the two subscripts stand for the two VI schedules. Since the reinforcement probabilities for the schedules differ on the basis of the product of the scheduled reinforcement rate and the time since the last response to the schedule, the momentary-maximizing rule reduces to

$$t_2 > t_1 \lambda_1 / \lambda_2. \qquad (4)$$

This inequality states that for concurrent VI VI, momentary maximizing requires that a response be made to the minority schedule (the schedule with the lower scheduled reinforcement rate; by convention choice No. 2) only when the time since a response to that schedule exceeds the time since the last majority choice (a response to the schedule offering more fre-

Table 1

Experimental Conditions for Concurrent VI VI

Condition	Bird #						Sessions
	CO96		CO123		CO104		
1	VI 60	VI 60	VI 60	VI 60	VI 60	VI 60	90
2	VI 180	VI 60	VI 180	VI 60	VI 180	VI 60	90
3	VI 60	VI 180	VI 60	VI 180	VI 60	VI 180	60
4	VI 60	VI 60	VI 60	VI 60	VI 60	VI 60	30
	CD129		CD117		CD148		
1	VI 180	VI 60	VI 180	VI 60			30
2	VI 60	VI 180	VI 60	VI 180	VI 60	VI 180	30
3	VI 60	VI 60	VI 60	VI 60	VI 60	VI 60	17
4	VI 240	VI 60	VI 240	VI 60	VI 240	VI 60	15
5	VI 90	VI 180	VI 90	VI 180	VI 90	VI 180	16
6			VI 180	VI 30	VI 180	VI 30	15

quent reinforcement) by a ratio greater than the ratio of scheduled VI reinforcement rates.

The Clock-Space Representation

The momentary-maximizing rule suggests that each choice during concurrent VI VI can be represented as a point in a two-dimensional clock space, the axes of which are the times since the previous response to each schedule (t_1 and t_2 in Equations 2 and 3). In the space, the momentary-maximizing criterion can be drawn as a switching line through the origin, of slope equal to the ratio of scheduled reinforcement rates. If the minority-choice time is represented on the ordinate, then behavior conforms to momentary maximizing when points representing minority choices fall between the switching line and the ordinate, and when majority choices fall between the switching line and the abscissa.

The actual position of points in the clock space for a given pattern of responses does not always agree with first intuition. The easiest way to make the representation clear is to generate the coordinates of a repeating sequence of choices and to observe the actual position of these points in the clock space. One such sequence appears in Figure 1 (borrowed from Staddon et al., 1981). The switching line for this clock space is of slope 3 indicating a 3 to 1 ratio of scheduled VI rates, e.g., a hypothetical VI-60 VI-180 schedule. If the animal makes a choice, say, every one second, then we would see the following sequence of coordinates. The first choice occurs by convention at time coordinates (1,1) and is for Response 1. At this point t_1 resets. The animal waits another second for its next choice, at which time the coordinates are (1,2). Since t_2 does not exceed t_1 by a ratio of 3, the slope of the switching line, Response 1 is again made. Again, t_1 resets and the animal waits a second for its next choice. After another second, the time coordinates are (1,3). This ambiguous situation, in which reinforcement probabilities are equal, will occur when choices are strictly periodic and the scheduled VI reinforcement rates form an integral ratio. For this example, we assume that reinforcement probability for the minority choice must exceed majority-reinforcement probability and thus Response 1 is again made. When the next choice is made the time coordinates are (1,4). This time Response 2 occurs and following thereafter a fixed set of coordi-

nates: Response 1 at (2,1), Response 1 at (1,2), Response 1 at (1,3), Response 2 at (1,4), and so on. We call this repeating sequence a momentary-maximizing trajectory.

As long as the animal makes choices at fixed time intervals, we can observe how some simple deviations from momentary maximizing appear in the clock space. For example, perseveration shows up as a line perpendicular to the appropriate time axis and extending beyond the switching line, e.g., (1,2), (1,3), (1,4), (1,5). On the other hand, if the animal switches prematurely to the minority response, this would truncate the triangle circumscribing the trajectory, e.g., (1,2), (1,3), (2,1), and so on. In the extreme, simple alternation appears as two
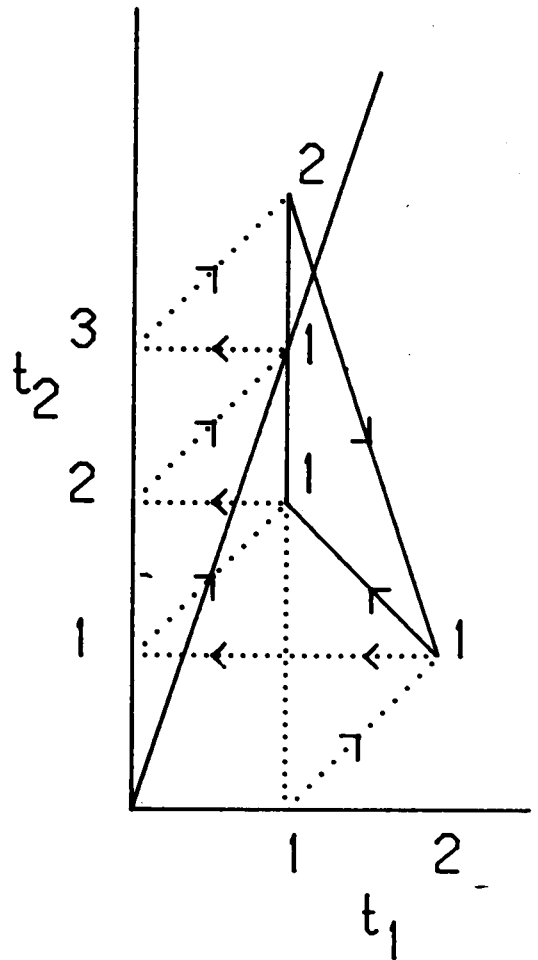

Fig. 1. Hypothetical clock-space trajectory for concurrent VI 60 VI 180. Dotted lines show how time since the last response resets with each choice made. The filled triangle shows the stable trajectory after initial choice sequence.

points, one for Response 1 at (2,1) and the other for Response 2 at (1,2). An increase in the choice frequency translates into a trajectory closer to the origin, whereas decreasing frequency displaces the trajectory away from the origin. These are the only variations in response patterns in the clock space if choices are made with fixed periodicity.

It is important to note that the switching-line analysis of momentary maximizing just establishes a criterion; it does not prescribe how that criterion shall be met. Equation 4 does not specify when a choice should be made, only which response should be made when a choice occurs. No particular pattern of responding, such as appears in Figure 1, is required by the analysis or assumed by it. Time between choices made by pigeons on concurrent VI VI is typically quite variable and, therefore, simple trajectories are not obtained. The most practical method we have found for

anticipating distributions of points within the clock space is to simulate responding with a specific choice rule.

We can expect to see two types of regularity in the clock space: (1) order due to periodicities in responding (i.e., any temporal regularity in choice or pattern of choices), and (2) order due to the location of responses with respect to the switching line. Figure 2 shows three types of idealized performances as a basis for comparison with actual behavior. To avoid confusion between the points for each choice, Response 1 choices are in the left panel, Response 2 choices in the right. Figure 2A shows simulated data in which responses occur randomly in time, but all choices obey the momentary maximizing rule for a hypothetical VI ratio of $\frac{1}{3}$, i.e., perfect momentary maximizing but no periodicity in responding. All responses are on the correct side of the switching line but are otherwise evenly distributed throughout the space.

Figure 2B displays simulated data where choices are also made randomly in time but biased to produce a ratio of response rates that matches the hypothetical $\frac{1}{3}$ ratio of VI values. In this case, there is neither periodicity in responding nor location with respect to the switching line, but the molar outcome of matching is preserved. Notice that matching by itself does not entail any organization with respect to the switching line: Responses are distributed uniformly throughout the space without regard to the switching line.

Figure 2C displays perfect momentary maximizing, with a 1/1 VI-schedule ratio, again with a random time between choices. In addition, this simulation illustrates the effect of a fixed, minimum time to change responses (changeover or switch time: CO). The diagonal of the clock space represents the points at which the times since each response are equal. Only when changeover time is zero can this condition be met. Since changeover time in practice will always diverge from zero, the changeover response must also diverge from the diagonal. A depopulated region of the space, because of switch time, is always along the diagonal, regardless of the location of the switching line and independent of conformity to momentary maximizing.

These sample simulations show that the clock-space representation does not create apparent order in the absence of some structure
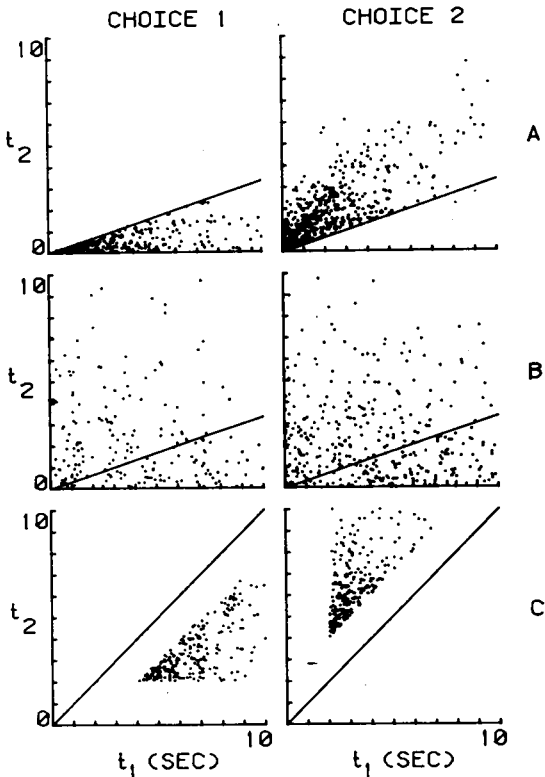

Fig. 2. Simulated data plotted in the clock space. Left and right columns are for two different choices. (A) Perfect choice on concurrent VI 180 VI 60 with random choice time. (B) Random choice on concurrent VI 180 VI 60, which results in matching. (C) Perfect choice on concurrent VI 60 VI 60 with minimum changeover (CO) time. Both time axes are in seconds.

in responding. Further, momentary maximizing does not require any pattern or periodicity, apart from dictating an area above or below the switching line in which respective responses must lie.

RESULTS

Pigeons' responding on concurrent VI VI is imperfectly periodic. Figure 3 shows two sets of representative data from one bird during training with concurrent VI 60 VI 60. Each

pair of panels displays all the responses made during a single session. As before, each response, 1 or 2, appears in a separate clock space to avoid confusion. The figure shows several characteristic features. First, points are not uniformly distributed in the space. On the whole, points are denser near the origin than away from it, implying that short interresponse times (IRTs) are more frequent. Second, thin bands of points often appear along a time axis —along the t_1 axis for Response 2 and along
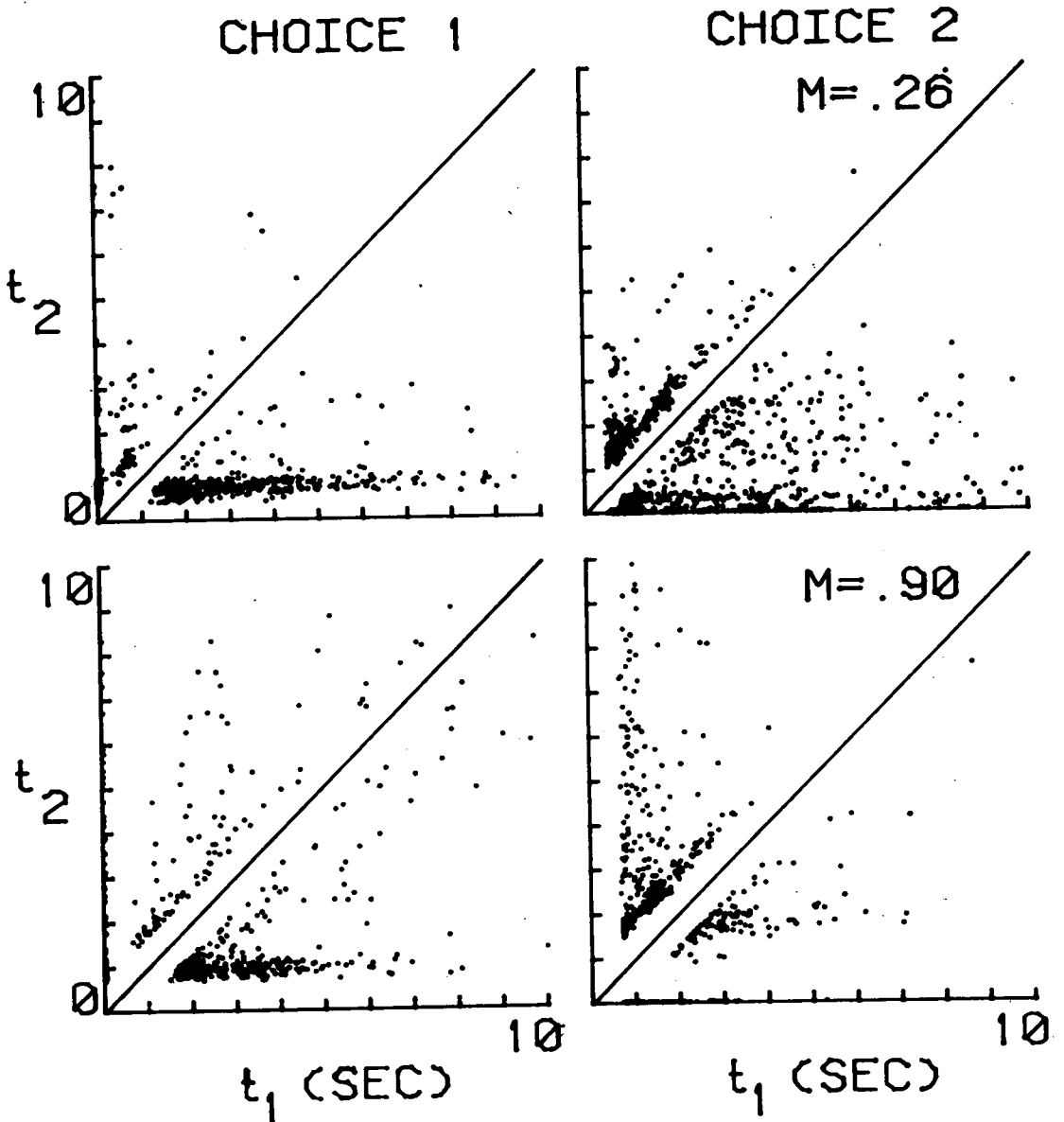

Fig. 3. Two sessions of data from concurrent VI 60 VI 60 sec for Bird CO123 represented in the clock space. (Top) Session 2, Condition 1. (Bottom) Session 24, Condition 1.

the t_2 axis for Response 1. Such bands, most prominent in Figure 3A, represent rapid "bursts" of key pecks, a common characteristic of pecking by pigeons (cf. Blough, 1963). The bands parallel to the t_1 axis in the right-hand panel of Figure 3A are bursts of successively greater number. A third prominent feature is the depopulated region along the diagonal and along the time axis of the response. This clear region has nothing to do with the switching line but represents a minimum CO time (cf. Figure 2C). Data from some sessions show points that lie very close to the diagonal. Thus, the relatively constant minimum CO observed in these and other data is a preferred switching rate rather than a mechanical limitation on the bird's ability to shift keys.

In addition to simple temporal properties of responding, we can also use the clock space to investigate how well behavior conforms to the switching-line criterion. The relative densities of points, the number and location of clusters, and the relative number of points in gross regions of the clock space can be roughly estimated. For example, relatively more points fall on the correct side of the switching line in Figure 3B than in Figure 3A, although for both data sets the greatest number of points is located near coordinates (1,2). By visual estimate, the conformity to momentary maximizing is better in Panel B (late in training) than in Panel A (early in training).

Figure 4 shows clock spaces for two pigeons, one on concurrent VI 180 VI 60 sec (top four panels) and one on VI 60 VI 180 (bottom four panels). The top four panels show two sessions: one early in training, when momentary maximizing is relatively poor, and one later in training, when momentary maximizing has improved. In Figure 4A (early) two dense regions appear for (majority) Response 2: a region between the switching line and the t_2 axis representing changeovers, and a region between the diagonal and the switching line. For (minority) Response 1 a dense, arrow-shaped set of points appears above and to the left of the line. Although most occurrences of Response 2, the majority response, meet the switching criterion, many occurrences of Response 1 indicate premature switching to the minority choice. A large change appears in Figure 4B (late). Two dense regions still appear for the majority response, more or less as before, but the minority choice is now represented primarily by a single

band of points beginning slightly above the switching line and extending through and below the switching line. The most dense regions lie near or just below the switching line, with density decreasing with increasing distance from the line. There is a small cluster of Response 2 bursts on the t_2 axis and another between the axis and the diagonal. The second group represents additional, perseverative responses after the switch from majority to minority choice. A smaller number of points can be seen distributed irregularly throughout the rest of the space. Figure 4B also shows the depopulated regions corresponding to minimum changeover time. Panels C and D in Figure 4 illustrate the typical consistency of momentary-maximizing performance across sessions. The data in these two pairs of panels are for one animal on successive sessions. The pattern is highly similar for these two days; this result is typical.

Despite small differences in detail, the global pattern of momentary maximizing is maintained for all animals in most sessions of all conditions. Majority choice responses occur in sequences of varying length, whereas minority choice is more often a single response, or short burst.

A Figure of Merit for Momentary Maximizing

The clock space reveals interesting temporal properties of behavior but does not allow for an easy assessment of the quality of momentary maximizing. Since choices are not perfectly periodic, we cannot expect to see a reliable pattern such as the trajectory of Figure 1. Further, a simple count of responses falling on the correct and incorrect side of the switching line provides only a crude estimate of the pigeon's adherence to the momentary-maximizing criterion.

If scheduled VI reinforcement rates are different, then reinforcement probability for the two responses changes over time at different rates determined by Equations 2 and 3. The momentary-maximizing criterion specifies that the response with higher reinforcement probability be chosen at any time. By any estimate, the larger the difference between the two reinforcement probabilities, the more severe the error if the pigeon makes the wrong choice. For example, a majority choice that appears in the clock space at a given distance on the
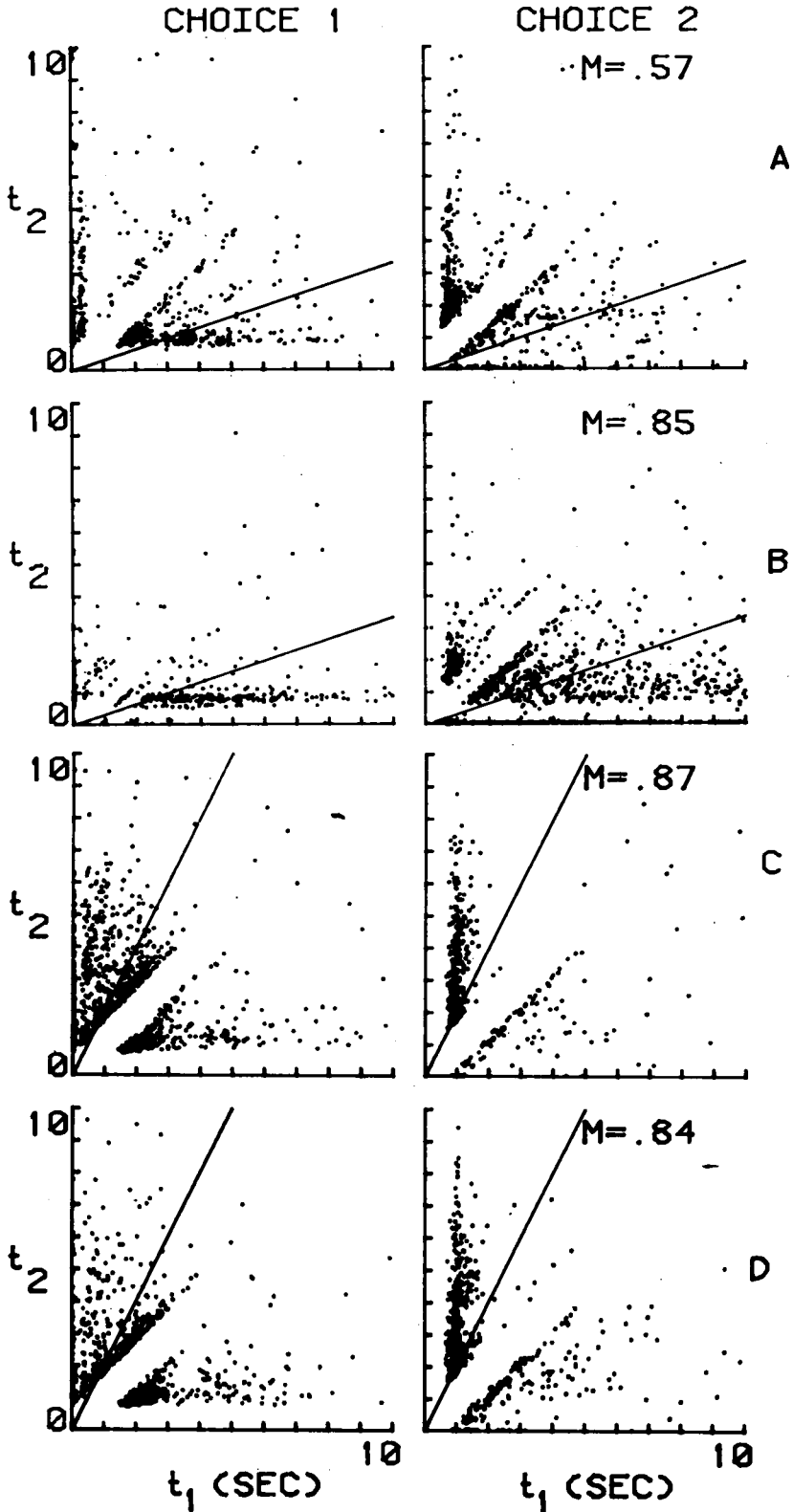
Fig. 4. Data represented in the clock space. (A) Session 2, Condition 2 for CO96. (B) Session 26, Condition 2 for CO96. (C) Session 12, Condition 5 for CD129. (D) Session 13, Condition 5 for CD129. All time axes are in seconds.

wrong side of the switching line (a perseverative error) represents a smaller reinforcement-probability difference, and therefore a less severe error, than a minority response at a comparable distance on the wrong side of the switching line; our measure should accommodate this difference.

Our estimate of the quality of momentary maximizing relies on the magnitude of reinforcement probability for each choice. When a choice is made, we calculate the reinforcement probability for each response using Equations 2 and 3. We then subtract the reinforcement probability of the response not chosen from the reinforcement probability for the response chosen. The difference is positive for correct responses and negative for incorrect responses, with the magnitude of the error corresponding to the magnitude of reinforcement-probability difference.

The procedure for obtaining the maximizing estimate is outlined in Table 2. The absolute value of probability difference associated with each choice is added into the appropriate cell of a 2 by 2 contingency table of response (1 or 2) by position with respect to the criterion (correct or incorrect). Each response represents a certain difference in reinforcement probability between the two schedules at the time of the choice. By computing the proportion of probability difference associated with correct responses (i.e., those obeying the momentary-maximizing criterion, cells a and d of the table) in the total probability difference (i.e., all cells of the table), we arrive at a general figure of merit for momentary maximizing, which we call m. The quantity m reflects the degree to which a response is sensitive to the maximizing criterion. Being a proportion, $m = 1.0$ if all responses are correct for maximizing (e.g., the simulation of perfect maximizing in Figure 1A), and $m = .0$ if all responses are incorrect. If responding is without regard

---

Table 2

Procedure for obtaining the momentary maximizing estimate m.

		Reinforcement Probability	
		$p(R\|t_1)>p(R\|t_2)$	$p(R\|t_1)<p(R\|t_2)$
	1	$\Sigma(p_1-p_2)=a$	$\Sigma(p_2-p_1)=b$
Response			
	2	$\Sigma(p_1-p_2)=c$	$\Sigma(p_2-p_1)=d$
		$m=(a+d)/(a+b+c+d)$	

---

to momentary maximizing (e.g., Figure 2B), then $m \simeq .5$.

Returning to the data sets displayed in the earlier clock spaces, we can now compare m values for sessions early (upper panels) and later (lower panels) in training. For Figure 2A m equals .26, while for 2B it is .90. Figures 3A and 3B represent m values of .57 and .85. Thus, obvious differences in the distribution of points in the clock space correspond to differences in the proportion of positive probability represented by our figure of merit m.

Table 3 provides a summary of m values for the first and last five sessions of each condition for each bird on concurrent VI VI. For Birds CO96 and CO123 the pattern of m changes is the same. The value for the first five sessions is lower than the last five, with one exception. Asymptotic levels of m are generally high, ranging from .71 to .90. Over conditions, these birds tend to adapt more quickly to changing schedule values. Bird CO104 performs well through the first three conditions, but thereafter the quality of momentary maximizing declines markedly.

Birds CD129, CD117, and CD148 show a pattern of changes across conditions similar to Birds CO96 and CO123. Each animal has a high m value in each condition over the last five sessions. For example, the range of asymptotic m values across conditions is from .77 to .80 for CD129, from .62 to .83 for CD117, and from .69 to .89 for CD148. In all cases, m is lower during the first five sessions, and in most cases substantially lower. On the whole, the quality of performance improves over conditions, although when scheduled VI rates are closer, improvement occurs more slowly.

Is It Really Hill-Climbing?

The observation that an animal always moves uphill is not necessarily evidence for a hill-climbing strategy. The animal could be on a broad slope so that any direction of movement within a 180° range would move it uphill. To be confident that the animal is following a hill-climbing strategy, we would need to show that any change in the direction of movement would reduce rate of ascent. Only in this way can we be sure that he is following the steepest-ascent rule. Thus, a high value of our summary statistic, m, is not sufficient by itself to prove a hill-climbing strategy. We need, in addition, to show that small changes in behav-

Table 3

Momentary Maximizing Estimate m for the First and Last Five Sessions in Each Condition

Bird:	Concurrent VI VI values											
	60/60		180/60		60/180		60/60					
	first	last	first	last	first	last	first	last				
CO96	.45	.71	.60	.81	.61	.81	.65	.83				
CO123	.27	.90	.83	.82	.67	.77	.46	.80				
CO104	.34	.62	.59	.64	.81	.63	.23	.32				
	180/60		60/180		60/60		240/60		90/180		30/180	
	first	last	first	last	first	last	first	last	first	last	first	last
CI129	.26	.82	.64	.83	.50	.77	.73	.85	.67	.83		
CI117	.42	.79	.34	.72	.18	.75	.78	.78	.48	.62	.78	.83
CI148			.57	.74	.32	.69	.61	.72	.62	.76	.73	.89

ior would reduce the value of m, i.e., to show that the existing pattern represents a maximum m value.

We describe two tests for hill-climbing. The first looks at the value of m obtained when a given data set is analyzed with respect to switching lines that deviate from the appropriate one. The second looks at the effect on m of displacing the sequence of responses to one choice positively or negatively with respect to the sequence of responses to the other choice. The questions in both cases are: Is there a shift for which m is a maximum? and is the maximum at, or close to, the scheduled conditions?

Shifted switching-line analysis. One straightforward way to examine how well behavior is adapted to the parameters of a particular schedule is to compare the obtained m value with other m values computed using the same data set, but with schedule values other than those actually used. In a sense, this amounts to shifting the switching line over a range and recording resultant changes in m. But, because m represents differences in reinforcement probability, we must compute m for particular VI values rather than changing the slope of the switching line per se. Specifically, we obtain several values of m for a data set by holding the majority schedule constant at VI 60 sec while varying the minority schedule from VI 60 sec to VI 600 sec in ten, 60-sec steps. The greatest m value over this range then indicates the schedule values that serve as a midpoint for a narrower range of ten minority-schedule values calculated from five, 10-sec steps on either side of the midpoint. The results of this shifted switching-line analysis will show: (1) if m is highest at the scheduled VI values, and (2) if the schedule value for which m is maxi-

mal (S_{max}) varies with changes in schedule values in a manner consistent with hill-climbing.

Figure 5 displays results of the shifted switching-line analysis with two types of simulated data. The upper panel shows m values obtained from random responding in perfect conformity with momentary maximizing (cf. Figure 2A). The lower panel shows random data, biased to produce matching between overall
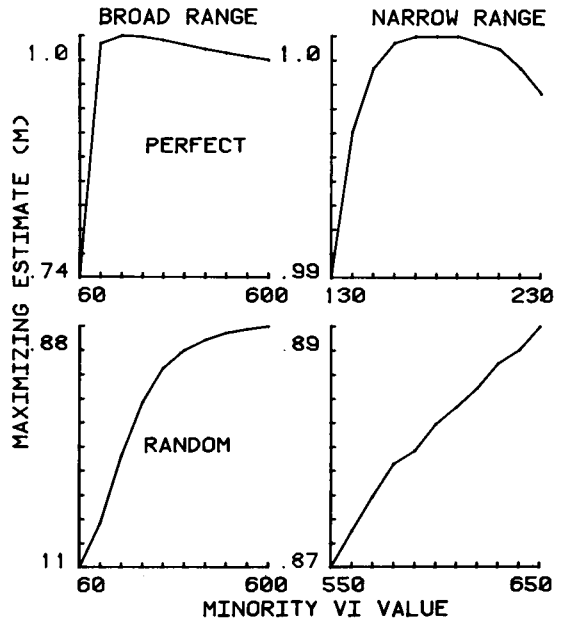

Fig. 5. Changes in m with shifts in the minority VI value for broad and narrow ranges of shift. (Lower) Simulated data with random choice time, which results in matching on concurrent VI 180 VI 60. (Upper) Simulated data with perfect momentary maximizing on concurrent VI 180 VI 60. The abscissa represents the shifted minority VI value. The ordinate represents the value of m at each minority VI value.

response and (scheduled) reinforcement rates (cf. Figure 2B). Left and right columns are for the broad- and narrow-range shift. The ordinate shows the range of m for each set. A maximum occurs at minority value VI 180 for the perfect momentary-maximizing simulation, as it should. On the other hand, there is no real maximum for the random simulation. The form of the function is continually increasing —because as the value of the hypothetical minority-schedule value decreases, the contribution of minority choices to the total proportion of reinforcement-probability differences becomes increasingly insignificant. (This property is not unique to m but is a feature of any measure that reflects differences, or ratios, of probabilities. See the General Discussion.) Behavior sensitive to the momentary-maximizing rule counteracts this effect and yields a maximum in the region of the actual schedule value.

Figure 6 shows representative data for one animal from two sessions of concurrent VI 60 VI 60. The lower panels show the results from a session early in training when $m = .24$. As with the random simulation of Figure 5, this session produces a curve that is constantly increasing, with no real maximum. The upper panels show a session late in training with $m = .90$. The curve increases smoothly toward the scheduled value of VI 60 in the broad range and in the narrow range shows a maximum at VI 70.

The same features are shown in Figure 7, which displays two sessions from a different animal on concurrent VI 180 VI 60. The lower panels show a session early in training in which m is very low. Again, the curve over both the broad and narrow range is continually increasing. The upper panels show a session later in training in the same condition. There is a clear maximum at the scheduled value of VI 180 for the broad range; the narrow range reveals that the best switching line is actually at VI 200.

In summary, when the value of m is high, the shifted switching-line analysis generates smooth curves with local maxima near the scheduled minority VI value. For sessions of low m, the function is increasing with no true maximum.

The most powerful test for hill-climbing is to show that the schedule proportion for which
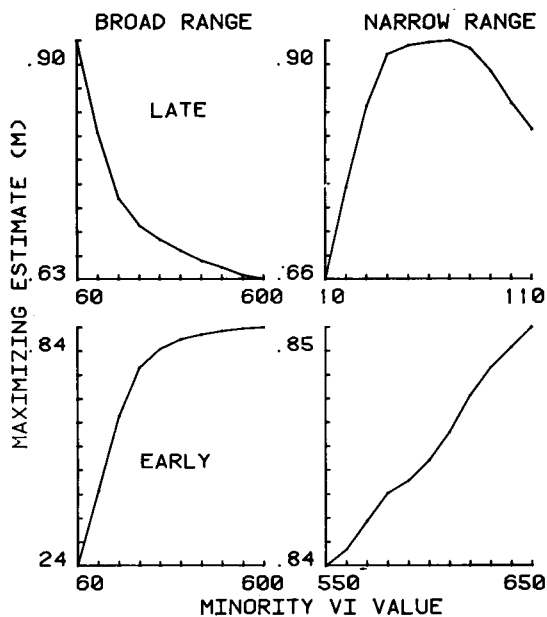

Fig. 6. Shifted switching line analysis for two sessions of concurrent VI 60 VI 60 for Bird CO123. (Bottom) Session 4, Condition 1. (Top) Session 24, Condition 1. The abscissa represents the shifted minority VI value. The ordinate represents the value of m at each minority VI value.
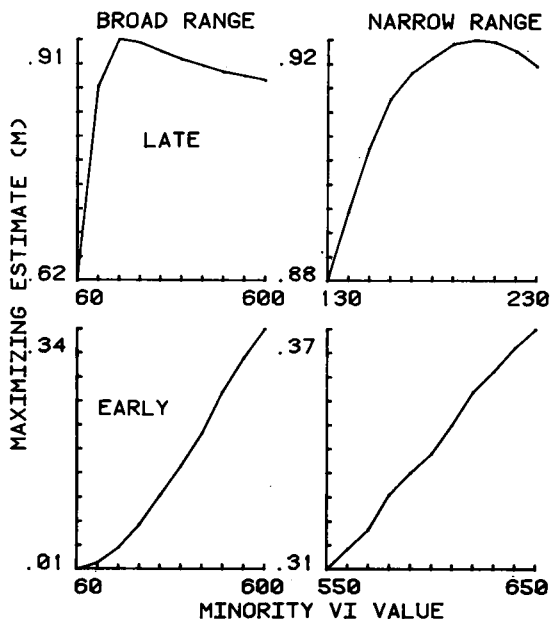
Fig. 7. Shifted switching-line analysis for two sessions of concurrent VI 180 VI 60 for Bird CD129. (Bottom) Session 2, Condition 1. (Top) Session 20, Condition 1. The abscissa represents the shifted minority VI value. The ordinate represents the value of m at each minority VI value.

m is a maximum, P_{max}, tracks imposed changes in scheduled proportion, P_{actual}. This is done in Figure 8, which shows the schedule reinforcement proportion ($a/[a + b]$, where a and b are scheduled rates of reinforcement) for which m is a maximum (by the shifted switching-line analysis) plotted as a function of the actual schedule proportion, for the asymptotic performance of all six pigeons in all the conditions to which they were exposed. Each point represents a single session; five asymptotic sessions are shown for each condition. For the most part, the points lie on or close to the 45° line representing perfect tracking. Bird CO104 was the only animal to fail to produce both consistently high m values over all conditions and appropriate values of P_{max}. For the other five birds, there is a close correspondence between P_{max} and P_{actual}. Surprisingly, perhaps, for every bird the greatest amount of variation in P_{max} occurs when both choices are equal (concurrent VI 60 VI 60). However, the range of variation over most values is small and centers on the actual value, and there is no other obvious trend in the display.

Shifted interresponse-time analysis. A second way to test for hill-climbing is to look at the effect on m of displacing the sequence of responses to one alternative with respect to the sequence of responses to the other.

The momentary-maximizing criterion specifies a precise relationship between the occurrence of either response and the times since both responses. Nonetheless, apparent sensitivity of behavior to the switching criterion could be due not to choice (i.e., a specific relation between responses to one alternative and responses to the other), but to a mixture of two independently controlled sets of events: Responses to each choice might be independent, with the distribution of responses in time for each determined only by its own generating distribution. The combination of the two interresponse (IRT) distributions might result, incidentally, in high m values, but this outcome would not reflect a choice strategy. Note that we use the term *IRT distribution* here in a more comprehensive sense than usual. Included is not only the frequency of different interresponse times (the usual definition) but also complete sequential information.

On the other hand, if the occurrence of a minority choice is affected by the momentary-maximizing rule, then every minority choice
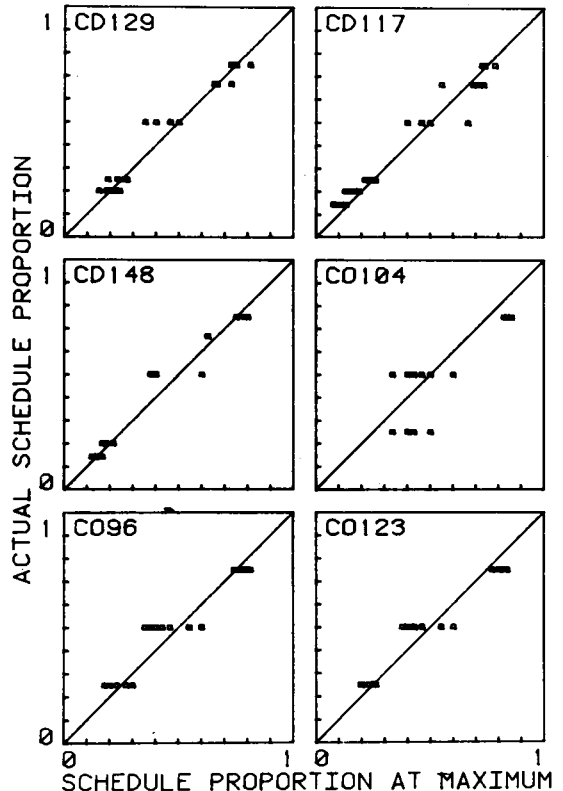

Fig. 8. Plots of the scheduled proportion of VI x to VI y versus the proportion of VI x VI y at the maximum value of m for all birds on concurrent VI VI. The points within each square represent the last five sessions of each condition that show a maximum. Graphs for each bird are as labeled.

time (t_2) should stand in an orderly temporal relation to the time since the last majority choice (t_1). Consequently, a shift in the distribution of responses to one choice relative to responses to the other should cause a drop in the computed value of m. Conversely, if the two distributions are independent, a shift should have no systematic effect. Since the separate IRT distributions are unaffected by these shifts, this analysis tests for the possibility that high m values are an accidental outcome of particular IRT distributions, independent of switches from one to the other. As before, we are interested in whether (over a small range of shifts) there is a maximum for m and whether the maximum is at zero shift.

Figure 9 shows results of the shifted-IRT-distribution analysis using sample simulated data. Each data set was shifted from −1 sec (minority distribution left-shifted) to +1 sec
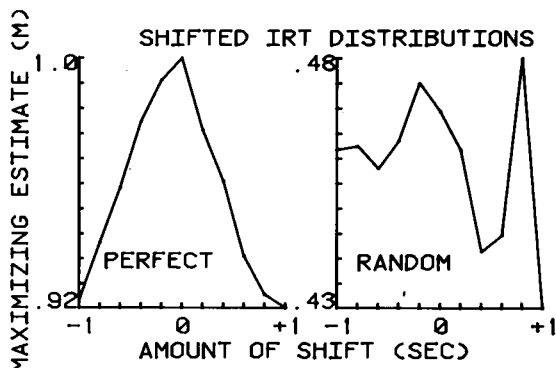
Fig. 9. Effects of shifting IRT distributions positively and negatively in time, with respect to one another. (Left) Simulated data for perfect choice with random choice time on concurrent VI 180 VI 60. (Right) Simulated data for random choice time with matching on concurrent VI 180 VI 60. The abscissa indicates the amount and direction (positive or negative) of the time shift. The ordinate indicates the *m* value corresponding to each shift.
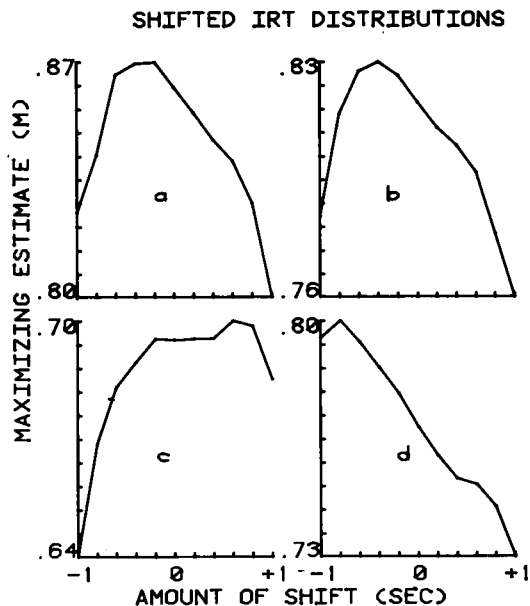

Fig. 10. Shifted IRT distribution analysis for four consecutive sessions of data from Bird CD117 on concurrent VI 180 VI 60. Panels are (a) Session 19, (b) Session 20, (c) Session 21, (d) Session 22, all for Condition 1. The abscissa indicates the amount and direction (positive or negative) of the time shift. The ordinate indicates the *m* value corresponding to each shift.

in five, .2-sec steps. The abscissa for each graph stands for the amount of time shift, and the ordinate gives the *m* value for each shift. The left-hand panel shows perfect momentary maximizing with choices occurring randomly in time (cf. Figure 4, top panels); the right-hand panel is for random responding biased to give molar matching (Figure 5, bottom panels). As expected, the perfect maximizing case on the left shows a local maximum at zero time shift. However, the random matching simulation on the right shows *m* varying in an erratic manner over a narrower range. The simulated data show that the shifted-IRT analysis provides another method of distinguishing a hill-climbing strategy from other reasons for a high *m* value.

Figure 10 shows shifted-IRT analysis of four consecutive sessions for a bird on concurrent VI 180 VI 60. These graphs show various degrees of adjustment to the schedule parameters, some maxima being nearer zero shift than others. For all of these sessions, *m* is reasonably high. The highest values of *m*, however, result in the most uniform and sharply peaked curves, such as those in Panels A and B. Although all the curves are slightly displaced from zero time shift, the amount of displacement is smallest for those sessions yielding highest *m*. In these data, maxima are usually somewhat to the left of zero shift, especially for sessions with high *m* values.

The data in Figure 10 are typical of good performance (high *m* values). Other data show that when *m* values are small, the shifted-IRT analysis shows no true maximum. Although the maximum is not always at zero time shift, it is not capricious. A session producing a high *m* score generally results in a time-shift function with a maximum in the region of zero. A slight displacement may mean an increase in *m*, with the shift typically toward the negative. On the other hand, sessions resulting in low *m* generally produce time-shift functions that increase monotonically to the left or the right over the range of shifts tried. Even on days of poor maximizing, responses to the two schedules are not independent. Rather, the animal seems to have misjudged either which is the majority choice, the most severe type of error, or how much more frequent reinforcement is for one schedule than for the other. The good approximation of momentary maximizing evident by other measures seems to reflect a hill-climbing strategy, not an incidental property of independent responding to the two alternatives.

EXPERIMENT 2
CONCURRENT VARIABLE-RATIO, VARIABLE-INTERVAL SCHEDULES

METHOD

Subjects

Four White Carneaux pigeons were used, two from Experiment 1, CR129 and CR117, and two other animals, CR196 and CR101, with prior experimental experience but none with concurrent schedules. All were maintained at 80% of their free-feeding weights throughout.

Apparatus

The experimental chamber and recording equipment were the same as in Experiment 1.

Procedure

The session duration, type of reinforcer, and reinforcer duration remained the same as in Experiment 1. The only procedural change from Experiment 1 was that one of the VI schedules was replaced by a constant-probability VR schedule. Table 4 shows the series of experimental conditions for the four birds.

Method of Data Analysis

A variable-ratio (VR) schedule delivers reinforcers on the basis of the previous number of responses. For a constant-probability VR schedule the distribution of ratio values is exponential. Since the reinforcement probability for any response on a VR is constant and equal to the reciprocal of the average ratio value M, the switching condition for momentary maximizing reduces to

$$t_2 > ln(M/(M-1))/\lambda_2, \qquad (5)$$

where by convention Response 1 is the VR alternative. Obviously, only the time since the VI-schedule response and the scheduled rates of reinforcement determine the switching rule. Since for a given schedule the values of λ_2 and M are constant, Equation 5 describes a switching line parallel to the t_1 axis (Staddon et al., 1981). In words, if a choice is made at a relatively short time since a VI response, then the response should be made to the VR. At longer times, the VI response should always be made.

RESULTS

Performance on concurrent VR VI usually follows the momentary-maximizing rule, although in general not as closely as concurrent VI VI performance. Figure 11 shows, for two birds, clock-space plots from early and late sessions of concurrent VR 15 VR 60 sec. Panels A and C are from poor momentary-maximizing sessions early in training; Panels B and D show better momentary maximizing from a later session for each bird. These plots share several common features with those for concurrent VI VI. First, responding on concurrent VR VI also shows a minimum CO time, evident from the absence of responses over an area along the diagonal and the time axis of the recorded response. Second, points are denser near the origin than away from it, indicating that shorter interresponse times again predominate. Third, response bursts also appear near each time axis.

Momentary maximizing in Panel A is poor, yielding an m of .07. Few responses are made to the VI schedule (Choice 2), while at least as many Choice 1 (VR schedule) responses appear above the switching line (incorrect responses) as below it. By the session in Panel B, many VI responses occur, appearing primarily as a thin band of points about half-way below the switching line and extending well above it. The VR responses are represented for this session by two arrow-shaped bands of points below the switching line and on either side of the diagonal. The m value is .86 for this session.

These results are typical. Although similar in several respects to performance on concurrent VI VI, the pattern of responding on concurrent VR VI differs most notably in the much shorter IRTs associated with the ratio schedule. Compare, for example, the "late" performances in Figures 11 and 3: Response 1 (the ratio response) in Figure 11 occurs at a

Table 4

Experimental Conditions for conc VR VI

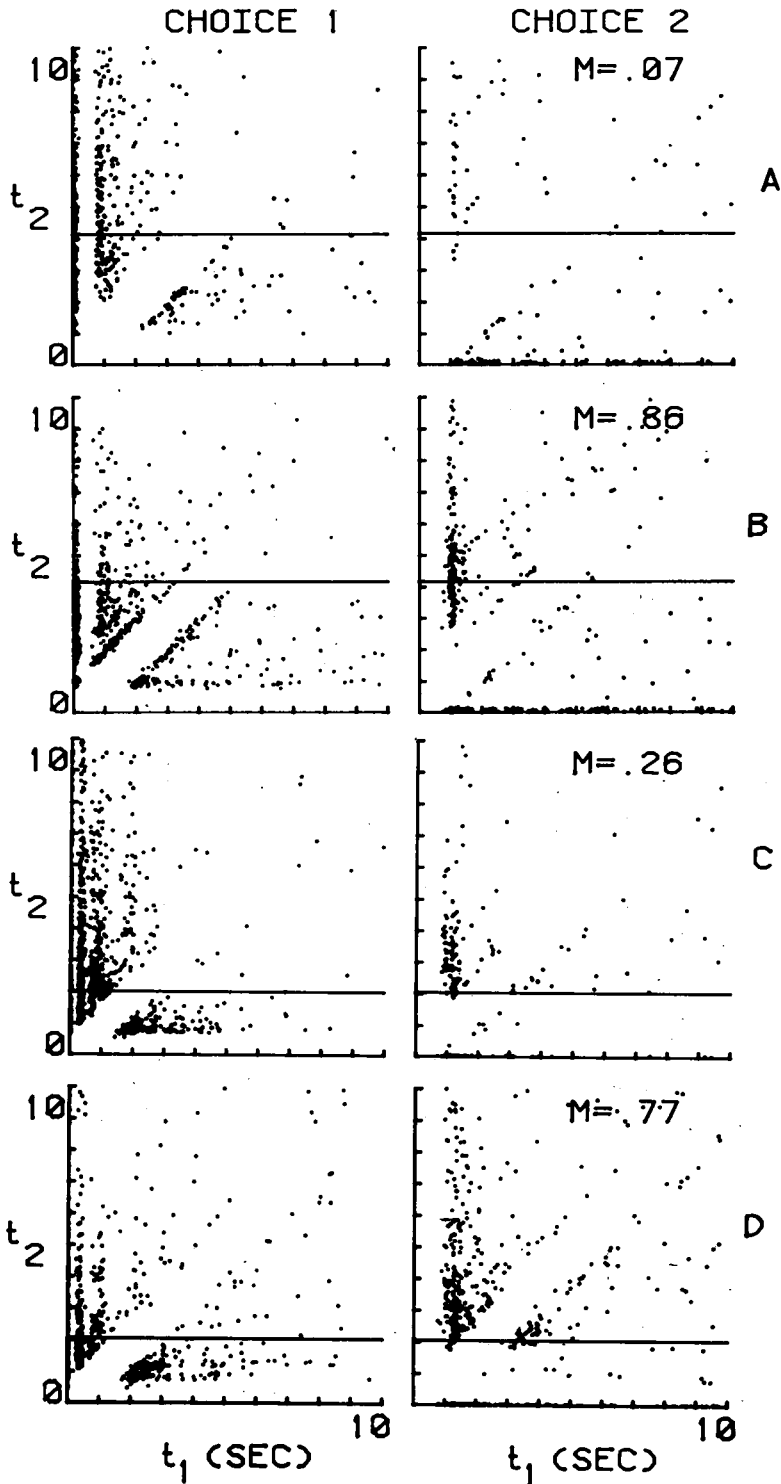| Condition	Schedule	Sessions	Schedule	Sessions
	CR129		CR117	
1	VR 30 VI 60	30	VR 30 VI 60	30
2	VR 60 VI 60	30	VR 60 VI 60	30
3	VR 15 VI 60	30	VR 15 VI 60	30
	CR101		CR196	
1	VR 15 VI 60	40	VR 60 VI 60	20
2	VR 60 VI 60	40	VR 60 VI 180	17
3	VR 30 VI 60	15	VR 60 VI 120	18

Fig. 11. Clock-space plots for representative sessions of concurrent VR VI. (A) Session 1, Condition 1 for CR101. (B) Session 5, Condition 1 for CR101. (C) Session 12, Condition 3 for CR196. (D) Session 15, Condition 3 for CR196. All time axes are in seconds.

Table 5

Momentary maximizing estimate m for the first and last five sessions in each condition.

Bird:	Concurrent VR VI values					
	30/60		60/60		15/60	
	first	last	first	last	first	last
CR129	.21	.30	.28	.87	.53	.65
CR117	.74	.70	.54	.94	.50	.84
CR101	.57	.85	.51	.76	.49	.88
	60/60		60/180		60/120	
	first	last	first	last	first	last
CR196	.53	.67	.65	.77	.39	.73

high local rate, indicated by bands parallel to that t_2 axis, whereas Response 2 (the interval response) shows no such bursts; both responses in Figure 3 show the interval, low rate pattern.

Table 5 displays m values for the first and last five sessions for each bird in each condition of concurrent VR VI. On the whole, performance is more variable for birds on concurrent VR VI, although average values are generally high at asymptote and lower during the first few sessions of a condition. CR117 and CR101 are the most consistent birds. For the last five sessions of each condition, the m value varies from .76 to .88 for CR101, while for CR117 the range is from .71 to .94. CR196 shows high m values for all conditions, though the absolute level is lower than for either CR101 or CR117. CR129 showed little improvement throughout Condition 1 but with continued experience over the remaining two conditions, performance improved to a point comparable with the other birds.

Quality of Momentary Maximizing on Concurrent VR VI

With minor adjustments, performance on concurrent VI VI and concurrent VR VI can be compared by employing the shifted switching-line analysis. By holding the VR value constant at the scheduled value and varying the hypothetical VI value, we can directly judge the quality of momentary maximizing. Figure 12 displays data from a single bird on concurrent VR 60 VI 180. The lower panel shows, as before, that a low m value (poor momentary maximizing) yields a continuously increasing function rather than a local maximum. The upper panel shows a local maximum at the scheduled value of VI 180. As with concurrent VI VI, the shifted switching-line function is

generally steeper in one direction. Over the range of schedule values we examined, this direction is toward the VR schedule.

As with concurrent VI VI, when maximizing is poor (indicated by a low m value for the session), the form of the shifted switching-line function is generally continually increasing. For these sessions, the only improvement in m occurs because incorrect VI responses make a less and less significant contribution to m as the switching line is shifted. Sessions producing a high m result in a function with a local maximum in the region of the scheduled VI value. During the best sessions, performance can be as close to optimal on concurrent VR VI as on concurrent VI VI. However, we found fewer sessions of high quality momentary maximizing on concurrent VR VI; and the switching-line location for which m is a maximum is not usually as close to the scheduled location as on concurrent VI VI. Figure 13 compares the VI schedule value at maximum to the actual VI schedule value for each bird over the last five sessions in each condition that showed a local maximum. In this figure, the abscissa represents the logarithm of the ratio (VI value at maximum m)/(actual VI value), and the ordi-
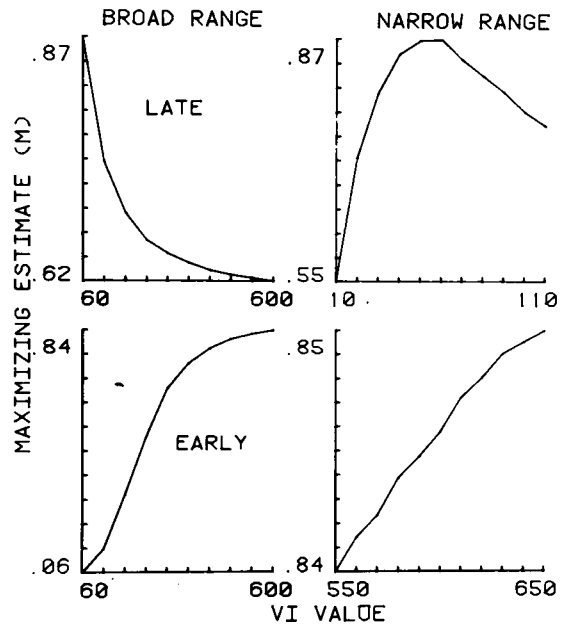

Fig. 12. Shifted switching-line analysis for two sessions of concurrent VR 60 VI 180 for Bird CR101. (Bottom) Session 1, Condition 1. (Top) Session 5, Condition 1. The abscissa represents the value of the shifted VI schedule. The ordinate represents the corresponding value of m for each shift.

CR101

CR117

CR196

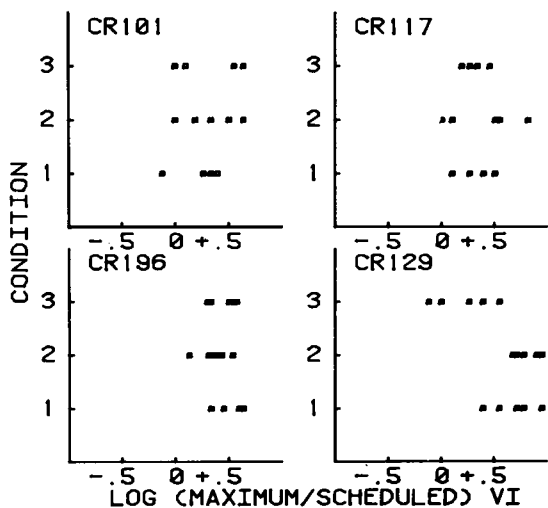CR129

LOG (MAXIMUM/SCHEDULED) VI

CONDITION

Fig. 13. Ratio of VI value at maximum *m* value over the last five sessions showing a maximum, over all conditions, for each bird on concurrent VR VI. The abscissa represents the logarithm of the ratio of the maximum to actual VI value. The ordinate displays the condition number.

nate shows the condition. On the whole, the quality of the estimate is poorer on concurrent VR VI (compare with Figure 8). The difference between the two values for most birds seems to be on the order of a factor of two. In terms of momentary maximizing, the birds are consistently underestimating the relative value of the VR.

The shifted-IRT-distribution analysis can also be applied to concurrent VR VI to separate properties of choice from the independent effects of the two schedules and to compare relative goodness of fit between concurrent VR VI and concurrent VI VI. For concurrent VR VI, we arbitrarily shifted the VI response distribution, since we cannot identify a minority choice. Looking over the same number of time shifts, we calculated the values of *m* using the scheduled reinforcement parameters at each shift. Figure 14 displays shifted-IRT distributions for four sets of data from a single animal. Because of the consistent error in estimation, the time range for the shift has been expanded from ±1 sec to ±2 sec. Figures 14 A and B show days of poor momentary maximizing. As before, these two functions increase in the negative time-shift direction throughout the range. In Figures 18C and 18D local maxima occur in the neighborhood of zero time shift.

As in the concurrent VI VI case, sessions pro-

ducing high *m* values yield a local maximum, and sessions producing a low *m* value do not—although the location of the maximum is further displaced from zero. Like concurrent VI VI, even on days of low *m* a consistent response pattern is being maintained, since the shifted-IRT function does not vary erratically. The strategy is merely poorly adjusted to the schedule parameters.

The lower quality of maximizing on concurrent VR VI is due primarily to an underestimation of *M*, the VR value. Dense parallel lines indicating response bursts are much more prominent in the VR side of clock-space plots and may help to explain the underestimation of the VR value. If each response burst is counted as a single unit, then the number of responses for each VR reinforcer will be underestimated, and the time to switch to the VI should be overestimated by the number of responses in the burst, two or three. Although it is not likely that the pigeon is counting responses in the same sense that humans count, it may be that whatever simple rule the bird is employing will reflect constraints in the animal's responding that do not (in the strict

SHIFTED IRT DISTRIBUTIONS

.09

.57

a

b

.01

.14

.82

.87

c

d

.03

.04

−2   0   +2 −2   0   +2
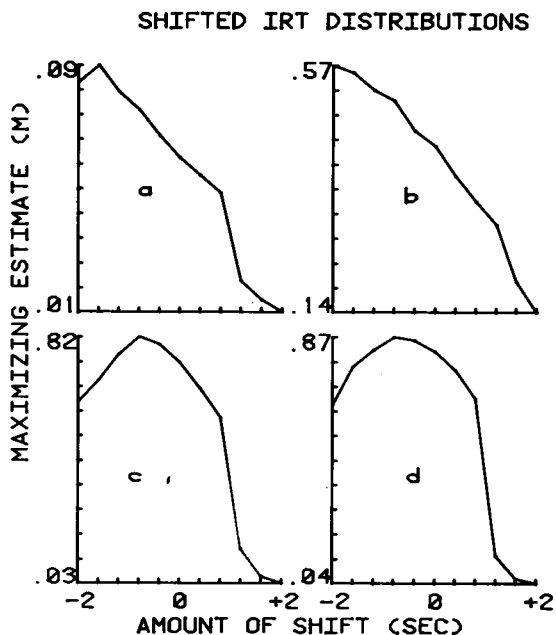AMOUNT OF SHIFT (SEC)

MAXIMIZING ESTIMATE (M)

Fig. 14. Shifted IRT distribution analysis for four successive sessions of concurrent VI 15 VI 60 for Bird CR101. Panels are (a) Session 4, (b) Session 5, (c) Session 6, (d) Session 7, all for Condition 1. The abscissa shows the amount and direction (positive or negative) of time shift. The ordinate shows the value of *m* corresponding to each shift.

sense) mean an inability to solve the maximizing problem. This argument is plausible because of the local maxima found with both the shifted switching-line and IRT-distribution analyses for sessions of high m. The form of the functions is much the same for concurrent VI VI and concurrent VR VI.

Because of differences between VR and VI schedule functions, a different response pattern should result on concurrent VI VI and concurrent VR VI if an animal obeys the momentary-maximizing strategy. For concurrent VI VI, the ratio of times since each response determines when a minority response should occur. On concurrent VR VI, absolute time since a VR response determines which response occurs. If choices are relatively infrequent with respect to the switching criterion, relatively more VI responses should occur; if choices occur more frequently, VR responses should occur more frequently. The expected relation between relative frequency of the two responses can be used as a test: Does the proportion of VI responses on concurrent VR VI increase as overall response rate decreases?

Table 6 displays the outcome of a product-moment correlation of number of VI responses and total responding for animals on concurrent VR VI. As we expect from momentary maximizing, all four birds on concurrent VR VI show substantial negative correlations, and three of these are highly significant: The more infrequently choices are made, the greater preference for the VI schedule. There is no systematic change in response proportion with changes in absolute rate for animals on concurrent VI VI.

GENERAL DISCUSSION

The questions with which we began this study are: On simple free-operant choice procedures, do pigeons consistently pick the alternative with the highest payoff probability (i.e., show momentary maximizing)? and does this represent a hill-climbing strategy? We answered these questions in a number of stages.

One preliminary task was to establish the formal criterion for momentary maximizing. This involved a graphic representation, the clock space, in which each choice is plotted. The switching line, which defines the momentary-maximizing criterion within the space, allowed a rough estimate of the distribution of correct and incorrect choices. Next we proposed a summary statistic, m, the proportion of positive probability difference, which provides a quantitative estimate of momentary maximizing. The proper evaluation of this statistic, which is important to the conclusions we draw, bears more detailed discussion.

In devising a quantitative measure of momentary maximizing, we have employed a number of alternative measures. For purposes of discussion, we shall consider the following: (1) number of correct responses; (2) distance from the switching line of each point in the clock space; (3) total distance from the switching line in a probability space; (4) ratio of reinforcement probabilities for each response; (5) proportion of positive-probability differences, the figure of merit m.

Obviously Measure (1), a simple tally of correct responses, is the crudest and least informative measure. Using Measure (2), we calculate the distance of each point from the switching line, counting correct responses as positive deviations and incorrect responses as negative deviations. The magnitude of positive and negative deviations then serves as an index of momentary maximizing. Measure (2) provides more precision but is flawed primarily because deviations of each point from the switching line in the time coordinates of the clock space do not accurately represent the probability differences with which the momentary-maximizing criterion is concerned.

Measures (3), (4), and (5) all concern differences in probability between responses when a choice is made. Using Measure (3), we would transform the time axes of the clock space into probability axes using the appropriate parameters of Equations 2 and 3. The diagonal in this probability space would represent the switching criterion (i.e., equal reinforcement prob-

Table 6

Product-moment correlation for proportion of VI responding and total responding during all conditions of concurrent VR VI.

	Bird	r	z	Significance level
VR VI				
	CR196	−.75	5.74	>.0001
	CR129	−.68	5.25	>.0001
	CR101	−.37	2.80	>.006
	CR117	−.16	1.25	>.2

ability). The magnitude of positive- and nega-
tive-probability deviations would then provide
an estimate of momentary maximizing. Mea-
sure (4) provides essentially the same index,
without generating a geometrical representa-
tion such as a probability space. If we take the
logarithm of the ratio of reinforcement-prob-
ability values for each choice, then a positive
value indicates momentary maximizing and a
negative value represents the opposite. Mea-
sure (5) is the figure of merit, m, discussed ear-
lier, which represents the proportion of posi-
tive-probability difference.

Measures (2), (3), and (4) have no upper and
lower bounds and are therefore more variable;
none indicates unambiguously the ideal that
should be approximated or standard against
which one is measuring. The total (or average)
amount of the deviation (or the log ratio) for
perfect behavior will depend upon the rule
that determines when a choice (not which re-
sponse) is made. If choice is perfect but choice
times are random, one value of these measures
will accrue. If choice is still perfect but there
is some periodicity in choice time, another
value will be obtained. One must, therefore,
assume a pattern of choice to properly evaluate
behavior using these measures.

We have settled on Measure (5), the figure
of merit m, for several reasons. It encompasses
the minimum necessary properties of a reason-
able statistic: (a) it measures probability differ-
erences; (b) larger positive values indicate bet-
ter performance, while larger negative values
represent poorer performance. As expected, we
find that Measures (3), (4), and (5) are highly
correlated. For example, Measures (3) and (4)
both give a positive session value for sessions
in which Measure (5) is above .50. Further,
Measures (3), (4), and (5) give similar functions
using the shifted switching-line analysis and
shifted-IRT analysis, although Measure (5) is
the most consistent across all of the analysis
techniques we have considered. Since Measure
(5) is bounded, it is less variable. Additionally,
Measure (5) has a common standard of value
1.0 for momentary maximizing regardless of
the pattern of choice times. Measure (5) in our
investigations has proved the most informative
measure of momentary maximizing.

Our summary statistic m reflects the mini-
mum properties for momentary maximizing
but does not make any additional assumptions

about perceptual or decision processes that fur-
ther constrain an animal's choice. For exam-
ple, it is likely that larger absolute values of
reinforcement probabilities require greater
proportional differences between alternatives
for accurate choice (i.e., a version of Weber's
law). However, inclusion of such an assump-
tion would blur the distinction between gen-
eral features of decision processes and the
momentary-maximizing criterion per se. The
momentary-maximizing criterion itself is in-
variant; through investigation we discover the
circumstances under which an animal meets
or fails to meet the criterion. This information
reveals what sorts of decision processes the ani-
mal must be employing.

Every statistic implies a theory of its appli-
cation. A high value of m alone cannot indi-
cate whether that value occurs due to: (1) the
sensitivity of behavior to changes in the sched-
ule, or (2) the insensitivity of the schedule-feed-
back properties to changes in parameters. We
used several converging measures to assess
conformity of behavior to the momentary-max-
imizing criterion. First, we showed that differ-
ent m values are correlated with different dis-
tributions of points in the clock space (e.g.,
Figure 3 and Table 5). We then carried out the
shifted switching-line analysis, which shows
how sensitive m is to the schedule parameters
and distinguishes between hill-climbing and
other sources of a high m score. High values
of m resulted in functions with a local maxi-
mum in the region of the actual schedule val-
ues (e.g., Figures 6 and 7). Next, we showed
that these local maxima occurred over many
different schedule values (see Figure 8). And
finally, we carried out the shifted-IRT-distri-
bution analysis, which provides the means to
distinguish how m depends upon choice strat-
egy or upon the simple combination of inde-
pendently controlled response distributions.
High m values resulted in functions with a
local maximum in the region of zero time shift
(e.g., Figure 9).

These tests turned out positive, often strik-
ingly so, over the majority of sessions in all
conditions. Birds on the concurrent variable-
interval, variable-interval procedure on the
whole showed better and more consistent per-
formance than birds on concurrent variable-
ratio, variable-interval. In both procedures,
pigeons reliably picked the higher-probability

alternative and did so in a way consistent with the hill-climbing hypothesis.

Correlates of Momentary Maximizing

The clock space shows a fairly consistent molar-response pattern, although none is required by momentary maximizing (i.e., choices may be random in time and still obey the rule). Birds on concurrent VI VI generally respond steadily to the majority VI schedule and occasionally sample the minority VI schedule with a single response or brief burst. This pattern was found by Silberberg et al. (1978) and our interchangeover time (ICT) distributions (not shown; cf. Hinson & Staddon, 1981) duplicate theirs. Once behavior has stabilized, the form of the ICT distribution varies little from good to bad momentary-maximizing sessions. The pattern of sampling the minority response and switching back to the majority response quickly in no way assures that the time of occurrence of the minority response relative to the majority response precisely obeys the momentary-maximizing criterion. Care must be taken in interpreting molar-response patterns as evidence for momentary maximizing, apart from the specific tests we have outlined above.

At first blush, momentary maximizing on concurrent VI VI seems to imply that the longer a response sequence to one alternative, the more likely the next response will be a changeover (e.g., Heyman, 1979). Although in the limit this must be true, variability in interresponse time allows changeover probability to be completely independent of runlength, whereas performance continues to conform to momentary maximizing. The relative independence of runlength and changeover probability is illustrated in Figure 15. The figure displays number of successive responses to each key (runlength) on the abscissa, and on the ordinate the conditional probability of a changeover for that runlength. The top two panels show data from sessions early and late in training, respectively, for one bird on concurrent VI 60 VI 180, and the lower pair show two sessions early and late in training for a different bird on concurrent VI 60 VI 60. In both cases, probability of changeover is largely independent of runlength. The major difference between early and late stages is the long runs of majority responses that sometimes occur early in training. These results are essentially the same as those found by Heyman (1979; with a
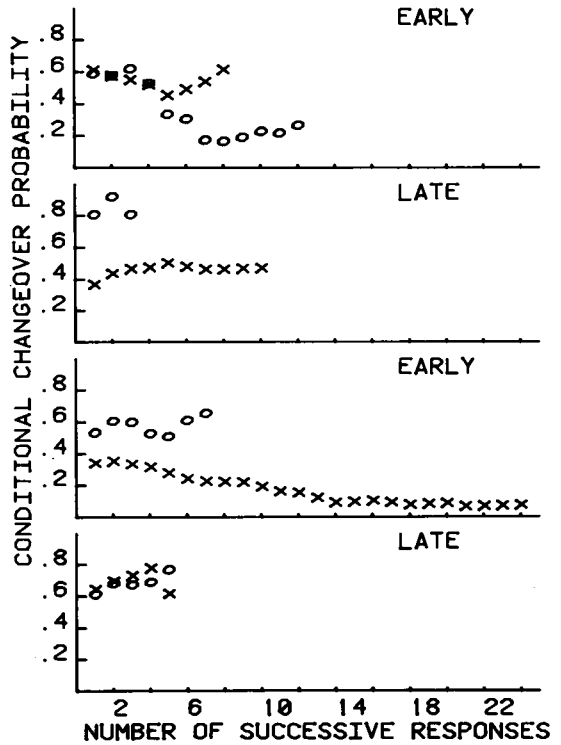

Fig. 15. Plots of probability of switching responses as a function of number of responses to an alternative. The top pair is for CO96 on concurrent VI 180 VI 60. (Early) Session 4, Condition 2. (Late) Session 26, Condition 2. The lower pair is for CO123 on concurrent VI 60 VI 60. (Early) Session 2, Condition 1. (Late) Session 24, Condition 1. The abscissa is number of consecutive responses to an alternative (runlength). The ordinate displays the conditional probability of a changeover at each runlength. Responses to Alternative 1 are represented as o's, and Alternative 2 responses appear as x's.

slightly different procedure) and taken by him as evidence against momentary maximizing. Since the *m* values for these asymptotic sessions were high, this conclusion is unwarranted. This same independence of runlength and changeover probability is also true of concurrent VI VR and even in simulations of perfect maximizing with random choice times (shown earlier in discussion of our analysis techniques). In none of these cases does conditional changeover probability increase with majority runlength, as previous studies have assumed it should.

In summary, although the molecular pattern of responding is consistent with momentary maximizing, we also find molar patterns of behavior consistent with previous studies. For concurrent VI VI schedules, the birds spend

longer and more variable amounts of time in responding to the majority schedule, whereas responding to the minority schedule takes a relatively shorter and less variable time. Minority runlengths are always much shorter than majority runlengths. For both concurrent VI VI and concurrent VR VI, the probability of changeover is reasonably stationary for all runlengths.

Deviations from Optimal Performance

The most common deviation from optimal performance in concurrent VI VI is underestimation of the minority-schedule value. Both the shifted-IRT and the shifted switching-line analyses often reveal a displaced maximum biased in favor of the majority schedule. As suggested earlier, this deviation may reflect the asymmetrical costs (in terms of "lost" probability) of early and late switches to the minority schedule. However, the shifted maxima may also be related to the minimum CO time characteristic of all sessions. The minimum CO may reflect what is called in other contexts "travel time" (Charnov, 1976; Krebs, 1978). The effect of travel time on the momentary-maximizing rule is somewhat different from its effect on rules for maximizing overall food rate —in patch foraging, for instance. For example, if changeover time is added to the time since a majority response for deciding when to switch, the bird will stay longer on the majority schedule. The CO time, however, does not affect switching back from the minority. Since the ratio of reinforcement rates favors the majority (by definition), a switch to the minority schedule should always be followed by a switch back to the majority. Thus, including the CO time as part of the maximizing estimate makes the majority schedule choice appear (in terms of the model) relatively richer.

On concurrent VR VI there is a more drastic overestimation of the VI value. In some ways, concurrent VR VI schedules present a simpler problem for momentary maximizing, because the animal need keep track only of the time since the last VI response (rather than the time since both responses). On the other hand, concurrent VI VI requires the application of only one process, timing, to both responses, whereas concurrent VR VI requires timing of the VI response plus a different process to estimate the VR value. This process need not be a strict averaging of the responses that produce

each reinforcer. If responding occurs at a near-constant rate, then the average time between reinforcers could be used as a derived measure of the VR value. The quality of maximizing under these conditions obviously depends on the variability in VR response rate.

One reason for the poorer momentary-maximizing performance on concurrent VR VI may be key-peck bursts—that is, different key-peck "weightings" (topographies) for the VI and VR choices. If response rate is being used to estimate the VR value, and bursts are treated as single, rather than multiple, responses, then the estimated VR value should always be smaller than the scheduled value. Failure to achieve optimal performance in this case may reflect a constraint in the response system rather than inability to maximize properly. A simple way to test this hypothesis would be to train animals on two types of concurrent VR VI schedules: (a) a schedule in which all VR responses contribute to completion of the ratio, and (b) a schedule in which short IRTs do not contribute to VR completion. The second schedule should result in performance closer to optimal for the present analysis. Another possibility is simply to present a brief stimulus after each response to eliminate bursts; this should also yield performances closer to the momentary-maximizing prediction.

During some sessions it is obvious that the birds are not maximizing well. Nonetheless, the global pattern of responding remains consistent. It seems that during sessions showing poor maximizing the birds are not obeying a different rule inconsistent with momentary maximizing. Rather, the choice pattern is simply poorly adjusted to the schedule parameters. This variation need not be interpreted as sheer avian ineptitude; without occasional variation in behavior, changes in response contingencies (in particular, in the schedule values) must go undetected. Optimal behavior requires some variation as a means of sampling for changes in the schedules—although framing this notion in a testable way poses something of a challenge.

A final source of deviations from perfect momentary maximizing must be mentioned. The theory cannot be true in general, simply because it predicts that the introduction of a changeover delay (COD) should completely abolish (rather than just diminish) switching

between schedules—because a COD means that switches are never immediately followed by reinforcement. Animals continue to switch, which implies that their time horizon must extend beyond the first postswitch response. Obviously the theory could be extended to deal with this, but at some cost in terms of simplicity; until this is done, however, we cannot be sure how much of the discrepancy we observe between actual and theoretical performance can be accounted for in this way.

What Is Optimal Behavior?

Optimal behavior, such as reinforcement-rate maximization, is often criticized on the basis that it cannot occur. The presumption is that something as clever as reinforcement-rate maximizing must be due to an equally sophisticated behavioral process—such as the molar-comparison strategy discussed in the Introduction (e.g., Herrnstein & Vaughan, 1980). However, few optimality theorists assume the literal calculation of reinforcement-rate functions that this argument implies. Instead, they suggest that a simple "rule of thumb," such as the hill-climbing process examined in this paper, has evolved which results in a good approximation to the ideal state. The apparent complexity of a behavioral adaptation in no way guarantees comparable complexity of the process underlying it.

Momentary maximizing is not the only kind of hill-climbing process that animals could use on concurrent schedules. An appealing alternative is that reinforcement-rate maximization might be the result of a simple choice rule that allows responses to track the occurrence of successive reinforcers. This strategy implies that a response to one alternative should be more likely in the period immediately following a reinforcement for that response. Reinforcement tracking of this sort, similar to the optimal giving-up time mechanism suggested by Krebs, Ryan, and Charnov (1974), can account for most molar features of behavior on concurrent VI VI (see also Staddon, 1980). Reinforcement tracking implies an easily detectable feature of behavior: a positive-recency effect of reinforcement (Menlove, 1975). (This may be the process Herrnstein and Vaughan [1980] term "melioration," although we cannot be sure since they do not define the term. See also Herrnstein, 1981.)
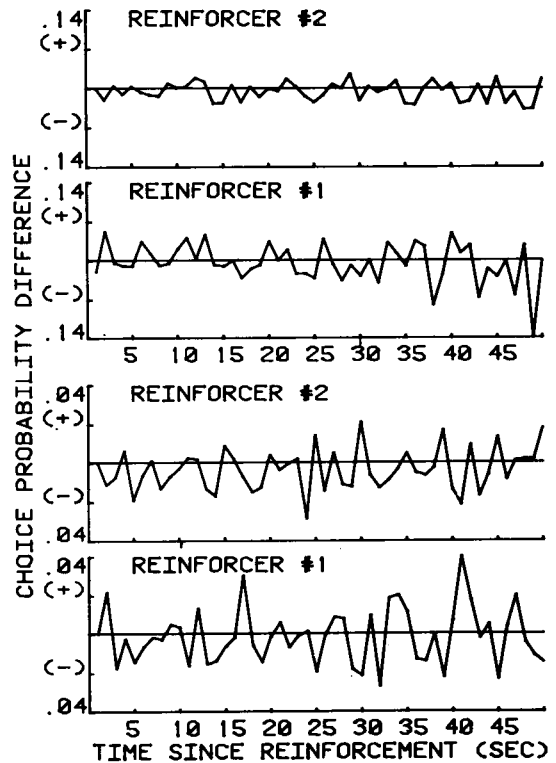

Fig. 16. Plots of the difference in conditional probability of either response (1 or 2), as a function of time since a reinforcer for either response. Each pair of graphs shows probability of responses after a reinforcer from Schedule 2 (upper) or Schedule 1 (lower). Two sessions on concurrent VI 60 VI 180 for CD117. (Top pair) Sessions 5, Condition 3. (Bottom pair) Session 20, Condition 2. The abscissa represents time since each reinforcer. The horizontal line in the middle of each graph represents zero probability difference between the two responses. Greater probability of the response to the same alternative as the reinforcer appears as a point above the horizontal line; greater probability of the alternative appears as a point below the line.

Figure 16 shows one of many tests for reinforcement tracking that we carried out. In this plot, responses are distributed into small time bins after each reinforcer. The upper and lower graphs of each pair show, for the two choices, a function relating the difference in the conditional probability of a response to the same schedule, or to the alternative, for each time bin (i.e., as a function of postfood time). A point above the horizontal middle line shows a net positive effect of a reinforcement on choice (same response more likely), a point below a negative effect. The use of a difference plot should exaggerate any recency effect, or

for that matter any immediate effect of reinforcement on subsequent choice. Only a handful of sessions out of hundreds (both VI VI and VI VR) tested show any positive or negative tracking effects. The vast majority are like those shown in Figure 16. The probability of a response to either choice is almost completely independent of postfood time. The pigeons did not track food deliveries.

The sensitivity of these animals to the momentary-maximizing criterion implies an adaptation to the schedule-feedback properties as programmed by the experimenter. Nonetheless, the schedule-feedback function is, in a sense, a fiction unless the animal responds in such a way that it can experience those feedback properties. At present, the means by which the birds are able to estimate the schedule values is unknown. The plots of S_{max} over conditions in Figures 8 and 13 clearly indicate the appropriate adaptation to a change in schedule values, but the variation in pattern during transitions between conditions defies simple description. Pigeons are able to measure times and weight them in the way required by momentary maximizing, but the means by which they do so remain to be discovered.

REFERENCES

Blough, D. S. Interresponse time as a function of continuous variables: A new method and some data. *Journal of the Experimental Analysis of Behavior,* 1963, **6**, 237-246.

Charnov, E. L. Optimal foraging: The marginal value theorem. *Theoretical Population Biology,* 1976, **9**, 129-136.

Fleshler, M., & Hoffman, H. S. A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior,* 1962, **5**, 529-530.

Herrnstein, R. J. A first law for behavioral analysis. *The Behavioral and Brain Sciences,* 1981, **4**, 392-395.

Herrnstein, R. J., & Loveland, D. H. Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior,* 1975, **24**, 107-116.

Herrnstein, R. J., & Vaughan, W. Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior.* New York: Academic Press, 1980.

Heyman, G. M. A Markov model description of changeover probabilities on concurrent variable-interval schedules. *Journal of the Experimental Analysis of Behavior,* 1979, **31**, 41-51.

Heyman, G. M., & Luce, R. D. Operant matching is not a logical consequence of maximizing reinforcement rate. *Animal Learning and Behavior,* 1979, **7**, 133-140.

Hinson, J. M., & Staddon, J. E. R. Maximizing on interval schedules. In C. M. Bradshaw (Ed.), *Recent developments in the quantification of steady-state operant behavior.* Amsterdam: Elsevier/North-Holland, 1981.

Krebs, J. R. Optimal foraging: Decision rules for predators. In J. R. Krebs & N. B. Davies (Eds.), *Behavioral ecology: An evolutionary approach.* Sutherland, Mass.: Sinauer, 1978.

Krebs, J. R., Ryan, J. C., & Charnov, E. L. Hunting by expectation or optimal foraging? A study of patch use by chickadees. *Animal Behavior,* 1974, **22**, 953-964.

Menlove, R. L. Local patterns of responding maintained by concurrent and multiple schedules. *Journal of the Experimental Analysis of Behavior,* 1975, **23**, 309-337.

Minsky, M. Steps toward artificial intelligence. *Proceedings of the Institutes of Radio Engineers,* 1961, **49**, 10-30.

Nevin, J. A. Overall matching versus momentary maximizing: Nevin (1969) revisited. *Journal of Experimental Psychology: Animal Behavior Processes,* 1979, **5**, 300-306.

Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. Economic demand theory and psychological studies of choice. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10). New York: Academic Press, 1976.

Shimp, C. P. Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior,* 1966, **9**, 443-455.

Shimp, C. P. Optimal behavior in free-operant experiments. *Psychological Review,* 1969, **76**, 97-112.

Silberberg, A., Hamilton, B., Ziriax, J. M., & Casey, J. The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes,* 1978, **4**, 368-398.

Staddon, J. E. R. *Limits to action: The allocation of individual behavior.* New York: Academic Press, 1980.

Staddon, J. E. R., Hinson, J. M., & Kram, R. Optimal choice. *Journal of the Experimental Analysis of Behavior,* 1981, **35**, 397-412.