JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1981, 36, 141-149 NUMBER 2 (SEPTEMBER)

MELIORATION, MATCHING, AND MAXIMIZATION

WILLIAM VAUGHAN, JR.

HARVARD UNIVERSITY

Pigeons were studied in an experiment involving two concurrently available response keys. Conditions were such that in the first condition the predictions of melioration (Herrnstein & Vaughan, 1980), minimization of deviation from matching, and maximization were identical: relative time on the right key should have fallen between .125 and .25, which in fact occurred. In the second condition, melioration predicted a shift in relative time on the right to between .75 and .875, which would involve a transient deviation from matching as well as a substantial drop in rate of reinforcement. All three birds eventually shifted their distribution of behavior to within the range predicted by melioration.

Key words: matching, concurrent schedule, maximization, key peck, pigeons

A number of experiments have demonstrated that, given a choice between two concurrently available variable-interval schedules (conc VI VI), animals and people tend, within some margin of error, to match both relative responses and relative time to relative obtained reinforcements (de Villiers, 1977; Herrnstein, 1961, 1970). That is, given two response alternatives, R and L, we find:

$$\frac{P_R}{P_R + P_L} = \frac{R_R}{R_R + R_L} \qquad (1)$$

$$\frac{T_R}{T_R + T_L} = \frac{R_R}{R_R + R_L} \qquad (2)$$

where P_R and P_L are responses to alternatives R and L, T_R and T_L are time spent at those alternatives, and R_R and R_L are obtained reinforcements on those alternatives. Equations 1 and 2 apply to asymptotic behavior; that is to say, distributions of behavior that appear not to be varying systematically.

This beachhead has in turn generated a vigorous controversy with regard to the mechanism, or process, responsible for matching at asymptote. One line of thought, deriving in part from neoclassical economic theory, holds that matching is the outcome of a maximization process. Shimp (1966, 1969), for example, has argued that, for conc VI VI, matching would be approximated if an animal always emitted that response with the highest momentary probability of reinforcement, which in turn implies that overall rate of reinforcement is maximized. Rachlin, Green, Kagel, and Battalio (1976) have shown, by means of a computer simulation, that matching is close to maximization of overall reinforcement rate, whereas Rachlin (1978) and Staddon and Motheral (1978) have argued that matching is analytically equivalent to maximization.

At the same time, there is some evidence against the maximization thesis. Herrnstein and Heyman (1979) have found matching, but not maximization, in a concurrent variable-interval variable-ratio (conc VI VR) experiment; Heyman and Luce (1979) have argued that on conc VI VI schedules, matching and maximization do not quite coincide, though their argument is not entirely persuasive (Rachlin, 1979). Finally, Herrnstein and Vaughan (1980) have reviewed several experiments which are inconsistent with maximization, but are consistent with a process they term melioration.

This latter process, which remains to be worked out with regard to a number of parameters, applies to situations in which several alternatives are being sampled. If the local rate of reinforcement on one alternative (i.e., R_i/T_i, number of reinforcements obtained from an alternative divided by time spent

---

This study was supported by NIMH grant MH-15494 to Harvard University. I thank R. J. Herrnstein for comments on an earlier draft. Reprints may be obtained from William Vaughan, Jr., Department of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138.

there) differs from the local rate on another alternative, then melioration says that the distribution of behavior will shift, in a relatively continuous manner, from the poorer to the richer alternative. The formula for melioration is given by:

$$R_D = \frac{R_R}{T_R} - \frac{R_L}{T_L} \tag{3}$$

in which R_D is the difference in local reinforcement rates from two alternatives, R and L. When R_D is positive, time spent at R (i.e., T_R) increases; when R_D is negative, T_L increases. Equilibrium is reached when $R_D = 0$ at which point Equation 3 implies Equation 2, the matching relationship for time spent responding.

Equation 3 is an idealization, in that no actual organism will respond to values of R_D sufficiently close to zero. For example, Herrnstein and Loveland (1975) find a closer approach to exclusive preference for the better of two VR schedules as the schedule values differ more; melioration, matching, and maximization all logically imply exclusive preference given unequal ratio requirements, no matter how small.

Although on some schedules (such as conc VR VR) melioration logically implies maximization of overall rate of reinforcement, in theory such a shift in behavior will occur whether or not overall rate of reinforcement increases (see Herrnstein and Vaughan, 1980). In fact, a decrease in overall rate of reinforcement can be consistent with melioration, given appropriate contingencies. The present experiment was designed to test the melioration hypothesis along these lines.

METHOD

Subjects

Three male White Carneaux pigeons with prior experimental histories were run. Just prior to the first condition described here, they had been exposed to special conc VI VI schedules for 26 sessions in the same chamber (results given in Figure 5.9 of Herrnstein and Vaughan, 1980). They were maintained at approximately 80% of their free-feeding weights.

Apparatus

A standard two-key pigeon chamber was employed. It was 33 cm high, 30 cm long, and 30 cm wide, with 2 keys centered on the front panel, their centers 22 cm from the floor and 10.5 cm apart. A standard grain magazine was used. The keys were transilluminated with white light, and the chamber was illuminated by two white Christmas tree bulbs. The keys required a force of about .14 N to operate, and a feedback click was provided. During reinforcement, which was access for 2.5 sec to mixed grain, only the hopper was illuminated, and pecks were not effective. The experiment was controlled by a PDP-8 computer.

Procedure

A response to either key initiated a timer which ran for two sec, unless reinforcement was delivered or the other key was pecked, initiating a timer there. Thus, if all interresponse times happened to be less than two sec, one of two timers would always run, except following reinforcement. Time on a key was accumulated while the associated timer was running; similarly, the VI tape (simulated by computer) associated with a key advanced only while the respective timer ran.

The rate at which each of the two VI tapes ran was a function of the distribution of behavior between the two keys, which was calculated as follows. After each four-minute period of responding, relative time on the right was defined as the proportion of that four-minute sample deriving from the right key. This proportion then set the rate at which the VI tapes would run for the next four minutes of responding. The specific functions relating local reinforcement rates on the two keys to relative time on the right are shown in Figure 1, for both the first condition (a) and second condition (b) of the present experiment.

Consider first the range of relative time on the right between 0 and .125. Given this distribution, whenever time accumulated on the right key, its VI tape advanced at a rate equivalent to a VI 20″ (180 reinforcements per hour); time spent on the left key advanced the other tape at a rate equivalent to a VI 60″ (60 reinforcements per hour). During responding on the right key, the left VI tape did not advance, and vice versa; while responding on neither key, neither tape advanced. When reinforcement set up, the VI tape stopped, and the next response to that key would in general produce reinforcement (during which neither timer ran). However, whenever a stopped

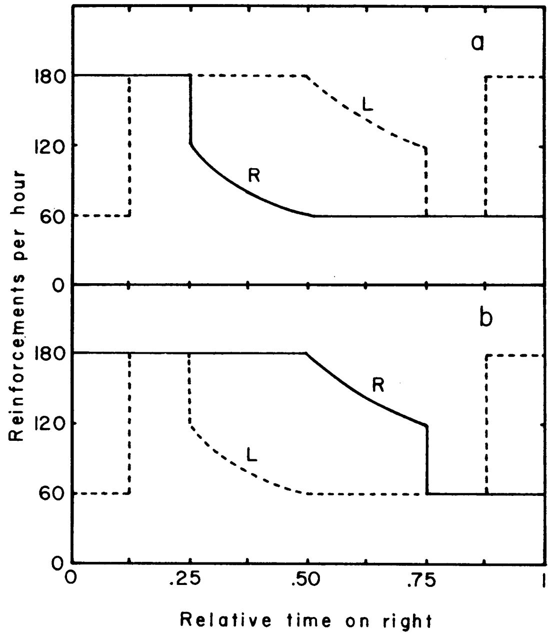

Fig. 1. Top: programmed local rates of reinforcement on left (L) and right (R), as a function of relative time on the right, during Condition a. Bottom: programmed local rates of reinforcement during Condition b.

timer was started, a 1-sec period had to elapse before a response could produce reinforcement, whether the last response had been to the same key or to the other key.

In the range between .125 and .25, each VI schedule paid off at 180 reinforcers per hour. Over the range of 0 to .25, the schedules did not differ between the first and second conditions; this also held for the range .75 to 1. Between .75 and .875, each key paid off at 60 reinforcers per hour, while from .875 to 1 the left key paid off at 180 reinforcers per hour while the right key paid off at 60 reinforcers per hour.

In Condition a, the left key (Figure 1, top, L) paid off at a higher rate than the right key (Figure 1, top, R) over the range of .25 to .75. In the second condition, these functions were simply reversed (Figure 1, bottom, R and L). Subject 2 required minor adjustments in local reinforcement rates for Condition b, as specified and discussed below.

During every 4-min period of responding, responses, time, and reinforcements for each key were recorded. In addition, total session time (including time responding and time not responding) and changeovers were recorded. Sessions were terminated as follows. After 1,800 sec of total session time or 35 reinforcers, whichever occurred first, the session teminated as soon as the next 4-min period of responding was over, unless 6 min elapsed without such a period ending, in which case the session terminated at that point. The initial reinforcement rates for each session were a function of the last 4-min period of responding the previous session.

RESULTS

Condition a was run for 27 sessions before the relative times for all pigeons appeared stable. For Pigeons 1 through 3, the averages of the relative times on the right for the last five sessions were .196, .160, and .148, respectively (see Figure 2). All but one of the individual sessions fell within the range of .125 to .25. Condition b was run for 56 sessions. Figure 2 shows, for each pigeon, the relative time on the right during all sessions of the experiment. The beginning of the first condition is indicated by a, and the beginning of the second condition is indicated by b. Subject 1 began to spend more time on the right on the first session of Condition b, and after seven sessions, had shifted to .78 relative time on the right. The average value for the last five sessions was .792.

For Pigeon 2, although relative time on the right was greater than .25 for some 4-min pe-

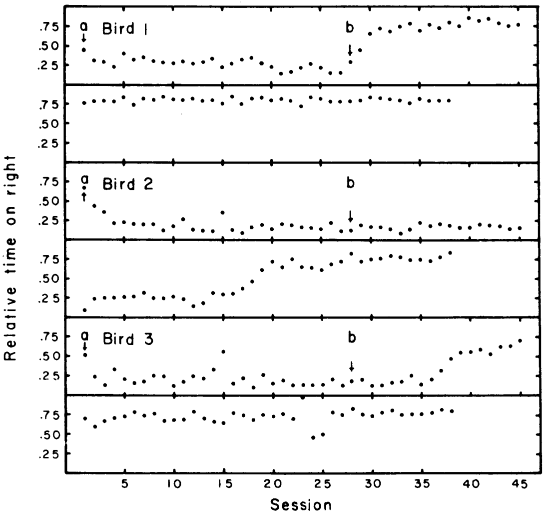

Fig. 2. Relative time on the right during Conditions a and b. Labeled session numbers refer to the upper panel for each bird. For lower panel session numbers add 45 to upper panel session numbers.

riods of responding, the difference in local reinforcement rates was apparently not sufficient (from the present point of view) to cause a significant shift in the distribution of time. Beginning with Session 46, the range over which the left key paid off at 60 reinforcers per hour was increased from 0 to .125 relative time on the right to 0 to .1875 (see Figure 3, left panel). This had the effect of increasing relative time on the right to some extent (Figure 2, Sessions 47 to 56, Subject 2). Finally, beginning with Session 57, reinforcement rate on the right key between .1875 and .25 was increased from 180 to 225 reinforcements per hour (Figure 3, right panel). During the last five sessions on this condition, relative time on the right was .768 for Subject 2.

Subject 3 began spending more time on the right beginning at about Session 37 of Condition b, with no adjustments of the parameters. The average for the last five sessions was .782. Thus, all three subjects eventually shifted their relative times on the right key from between .125 and .25 to between .75 and .875.

As the subjects shifted more time to the right during Condition b, they transiently increased deviations from matching and reduced overall rate of reinforcement, in accord with the experimental design. Figure 4 shows absolute deviation from matching, as given by the expression:

$$\left| \frac{R_R}{R_R + R_L} - \frac{T_R}{T_R + T_L} \right| \tag{4}$$

The abscissa is the overall rate of reinforcement, calculated over the complete session duration (excluding only feeder durations). The points labeled a show the average overall rate of reinforcement and the average absolute deviation from matching for the last five sessions of Condition a. The next point, labeled b′, plots those values for the five consecutive sessions during which relative time appeared to be changing most rapidly (see Figure 2). For Pigeon 1 this derives from Sessions 28 to 32; for Pigeon 2, Sessions 61 to 65; for Pigeon 3, Sessions 36 to 40. It can be seen that in the transition from Condition a to b, the average absolute deviation from matching increased, whereas overall rate of reinforcement decreased. The last point, labeled b, derives from the last five sessions of Condition b, by which time the absolute deviation from matching had dropped back down, whereas for all three birds overall rate of reinforcement had dropped by about 40%.

Table 1 shows the sums of responses, time measures, obtained reinforcers, and changeovers for the sessions plotted in Figure 4.

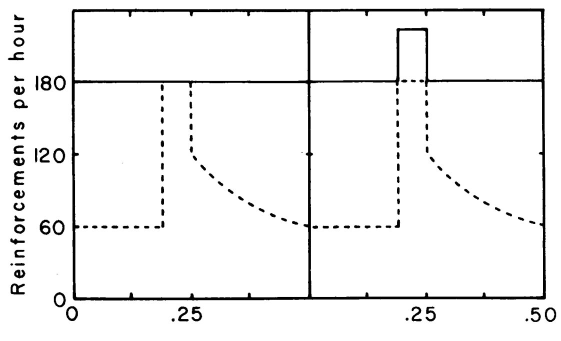

Fig. 3. Modified programmed local rates of reinforcement for Bird 2 as a function of relative time on the right between 0 and .5. Left: first modification of Condition b schedules. Right: second modification.

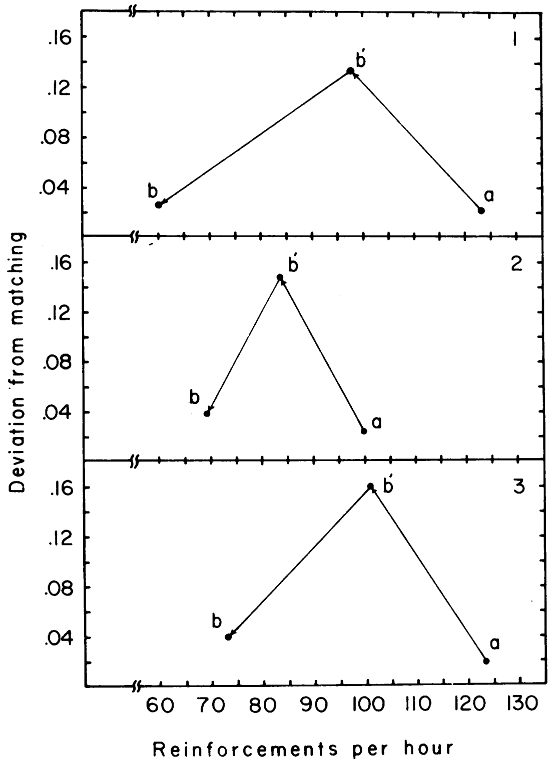

Fig. 4. Absolute deviation from matching and overall rate of reinforcement (calculated over the whole session, excluding reinforcer time) for the last five sessions of Condition a (a), while behavior was changing in Condition b (b′), and during the last five sessions of Condition b (b).

Table 1

Data in the form of sums for the last five sessions of Condition a, the five days during which behavior was changing most rapidly (b′), and the last five sessions of Condition b.

	Responses		*Time (Seconds)*			*Reinforcers*		
Condition	*Left*	*Right*	*Left*	*Right*	*Total*	*Left*	*Right*	*CO's*
a								
Bird 1	4920	1946	4262.7	1017.3	6308	173	37	638
Bird 2	9362	2115	4450.0	824.4	7333	167	36	291
Bird 3	5504	1203	4321.0	719.0	5808	163	32	401
b′								
Bird 1	3212	5730	2673.3	3326.0	7151	62	132	1384
Bird 2	4571	6391	2615.9	2664.1	8244	64	120	482
Bird 3	4181	3424	3374.5	2385.5	6562	77	107	382
b								
Bird 1	2337	5592	1638.2	6281.8	9246	29	124	753
Bird 2	3363	10696	1688.8	5751.2	9052	33	136	603
Bird 3	2056	6589	1768.7	6391.3	9038	32	151	419

(Since Figure 4 derives from averages across individual sessions, there is not an exact correspondence between Table 1 and Figure 4.) Although it might be expected that the proportion of time spent responding would decrease as reinforcement rate went down, there is in fact a slight increase from Condition a to Condition b. Bird 1's overall response rate showed some decline, but the other two birds remained approximately constant between Condition a and b.

DISCUSSION

The logic of the procedure arises in three competing theories of behavioral allocation, namely, melioration, maximization, and matching. Each theory hypothesizes a dynamic process leading to matching (at least under some conditions, variously delimited in different theories) in the steady state. For melioration (Herrnstein & Vaughan, 1980) the dynamic process acts to shift relative time from locally poorer to locally richer situations; for global maximization (Rachlin et al., 1976), it tends to maximize total reinforcement (or utility) across alternatives; for matching, it tends to minimize deviations from matching.

With regard to maximization, this experiment addresses theories which assume that animals behave so as to maximize rate of reinforcement as measured, say, over several sessions (cf. Rachlin, 1978; Staddon & Motheral, 1978). The experiment does not explicitly make contact with momentary maximization (Shimp, 1966, 1969), except as follows: if momentary maximization can account for the present results, then global maximization is not a logical consequence of that theory.

As a dynamic process, global maximization requires that an animal sample at least two distributions of behavior between the two keys. Given a single distribution, regardless of the outcome, there is not sufficient information to say in what direction behavior should shift so as to increase overall rate of reinforcement.

The same limitation applies to minimization of deviation from matching, as construed here. Again, given a single distribution of behavior, there is not sufficient information to say in which direction (if either) behavior must shift in order to reduce deviation from matching. The dynamic matching process was discounted as a possibility in Herrnstein and Loveland's study of conc VR VR (1975) and offered speculatively by Rachlin (1973), who suggested that matching may come about because animals find unequal local rates of reinforcement aversive, but does not seem to have any active adherents at present.

Each of the theories can be illustrated for the present procedure. Figure 5 shows, for Conditions a and b, overall rate of reinforcement as a function of relative time on the right. For Subjects 1 and 3, 180 reinforcers per hour (of time spent responding) would be received for spending between .125 and .25 relative time on the right, whereas, for all subjects, 60 reinforcers per hour would derive from spending between .75 and .875 relative time on the right. The dashed lines show the adjusted

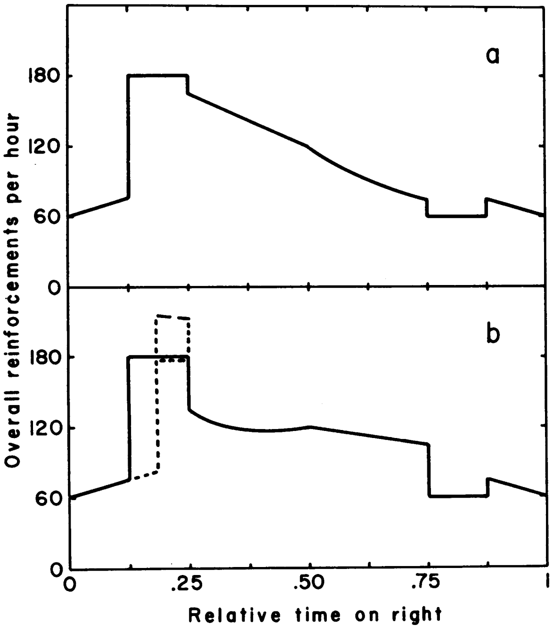

Fig. 5. Programmed overall rate of reinforcement, as a function of relative time on the right, for Conditions a and b. This figure may be derived from Figure 1. Reinforcement rate and relative time are both calculated on the basis of time spent responding. Dashed lines represent modifications for Bird 2.

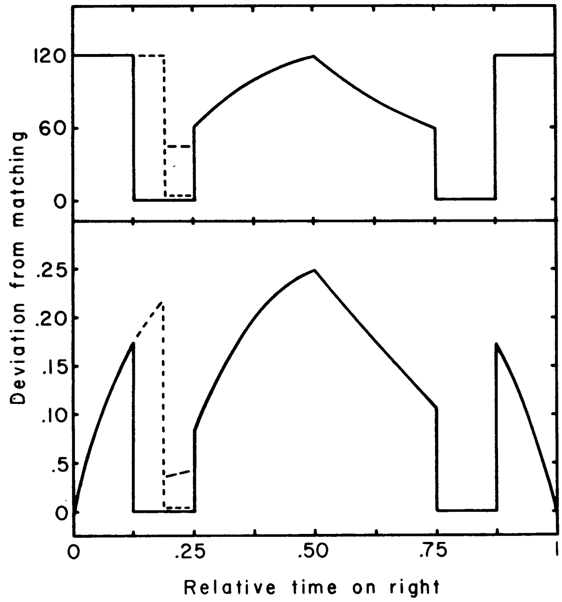

Fig. 6. Top: programmed deviation from matching, measured in terms of the absolute differences between local rates of reinforcement, as a function of relative time on the right. Bottom: programmed deviation from matching in terms of the absolute differences between relative time on the right and relative programmed reinforcements on the right. Dashed lines represent modifications for Bird 2.

values for Subject 2 in Condition b. The implication from a maximization point of view, then, is that relative time on the right should fall between .125 and .25 in both conditions (slightly altered for Subject 2 in Condition b). Behavior between .75 and .875 would be least in accord with a maximization analysis for all subjects. Yet, all subjects spent between .75 and .875 of the time on the right (averaging .78) in Condition b.

Figure 6 (top) shows deviation from matching measured as the absolute difference in local reinforcement rates, as a function of relative time on the right. The function is the same for Conditions a and b; dashed lines show adjusted values for Subject 2 in Condition b. Figure 6 (bottom) shows deviation from matching (for Condition b) measured as the absolute difference between relative time on the right and relative programmed reinforcements on the right; for Condition a the function is approximately the same. In both Conditions a and b, matching occurs between .125 and .25 and between .75 and .875. Hence, if a pigeon found itself in either of these troughs it should remain there, given that it behaves so as to minimize deviation from matching (measured either way). In fact, all three pigeons concluded Condition a in the left trough, but then shifted to the right trough in Condition b, climbing across greater deviations from matching (see Figure 4) as they did so. [Notice that Figure 6 (bottom) employs the same ordinate units as Figure 4.] At least for Subjects 1 and 3, we can reject minimization of deviation from matching as the dynamic process.

The analysis of the schedules in terms of melioration is shown in Figure 7. This shows local rate of reinforcement on the right minus local rate of reinforcement on the left, or net reinforcement rate on the right, as a function of relative time on the right. Within the range of 0 to .125 relative time on the right, the right side pays off at 120 more reinforcers per hour than the left, and hence melioration predicts a shift to the right (see Equation 3). The opposite contingencies hold between .875 and 1, and so by the same logic behavior should shift left. Between .125 and .25, and .75 and .875, the two local rates are equal, implying no net tendency to shift.

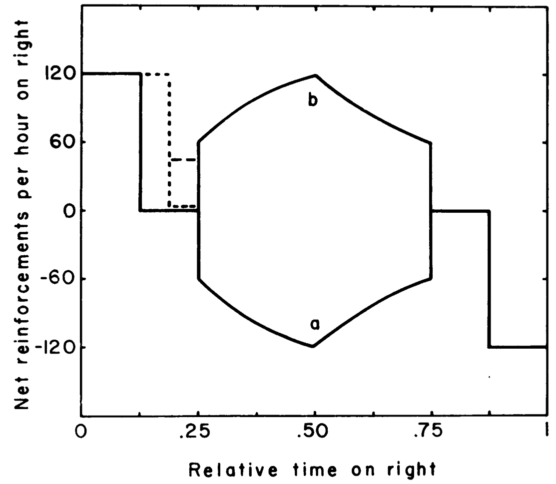

Fig. 7. Signed differences between programmed local rates of reinforcement on the right and on the left, as a function of relative time on the right, for Conditions a and b. This figure may be derived from Figure 1. Dashed lines represent modifications for Bird 2.

In Condition a, between .25 and .75, the left side pays off at a higher rate than the right. Within this range, then, behavior should shift left. In Condition b, these contingencies are reversed, and so behavior should shift to the right, until it falls within the range .75 to .875. Subject 2's Condition b procedure is shown as before by the dashed lines.

The general design of the experiment, then, involved, in Condition a, getting behavior to stabilize within the range .125 to .25. Given a variety of initial conditions, behavior within this interval would be consistent with maximization, minimization of deviations from matching, and melioration. In Condition b, according to the first two analyses, behavior should remain with the interval .125 to .25. According to melioration, if by chance behavior shifts sufficiently to the right of .25 for the bird to be affected by the unequal local rates of reinforcement, it should then move to the range .75 to .875, in spite of the fact that this would involve a transient increase in deviation from matching, as well as a sizeable reduction in overall rate of reinforcement. All three subjects unambiguously confirmed melioration as the dynamic process, even the slighly aberrant Pigeon 2, as can readily be explained.

Melioration permits a subject to remain unaffected by the transition from Condition a to b if it fails to sample significantly outside the interval between .125 and .25, within which local reinforcement rates remained equal. The day-to-day record for Subject 2 (Figure 2) shows stable allocations of time spent on the right. When the local rates of reinforcement were made unequal for its allocation, melioration then implies the observed shift in its allocation. The corollary is that a certain amount of behavioral variability protects an animal from missing changed circumstances, illustrated by Pigeon 2.

Since the present procedure is somewhat atypical, it may be appropriate to consider its more unusual features explicitly. In particular we should consider the continual adjustment of local reinforcement rate by successive 4-min samples of responding. It may be objected that such rapidly changing contingencies of reinforcement would not control responding. However, the smooth and generally rapid transition in performance from Condition a to b shows that responding was under quite good control. Moreover, Figure 8 plots the final levels of allocation for each subject in each condition, using the logged ratios of local time (T_R/T_L) and pecks (P_R/P_L) as a function of numbers of reinforcements (R_R/R_L) (see Baum, 1974). Considering that the procedure lacked a genuine COD, the observed amount of undermatching would not be excessive even for conventional conc VI schedules. Matching in the steady state, as well as transitions in the distribution of behavior from locally poorer to locally richer situations, may each be viewed as outcomes of the mechanism of melioration.

In their discussion of behavior on concurrent interval and concurrent ratio schedules, Fantino and Logan (1979, pp. 215-216) make the apparently innocuous argument that concurrent interval schedules sustain behavior on both keys because a higher rate of reinforcement is thereby produced. The present results, if taken at face value, imply that their argument is neither innocuous nor correct: rather than overall rate of reinforcement, it is local rate of reinforcement that governs the distribution of behavior.

When arguing against the possibility of a science of behavior, it is often choice (or "genuine" choice) that is invoked. A maximization analysis, if true, would defuse the above strategy by showing choice behavior to derive from response strength: that distribution of responses which tends to maximize reinforce-

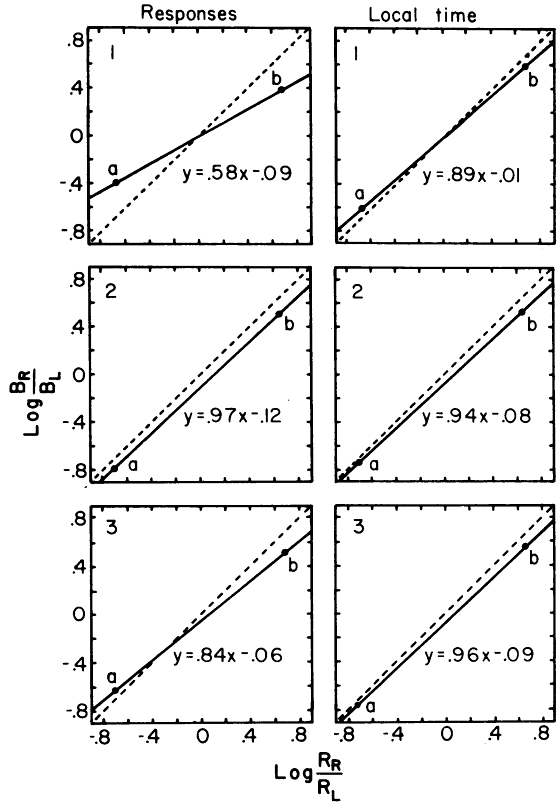

Fig. 8. Logged ratios of behavior as a function of logged ratios of obtained reinforcements. Left: response ratios. Right: time ratios, calculated on the basis of time spent responding.

ment rate would be most strengthened. It is also possible to deduce melioration from a response-strength analysis, as follows.

Suppose a pigeon receives food in the presence of two stimuli, but at different rates. By means of classical conditioning (cf. Rescorla & Wagner, 1972), that stimulus associated with the higher density of reinforcement will come to have higher value. Consider a response that causes a transition from that stimulus to the other. Since it leads from a situation of higher value to one of lower value, it should be punished (Baum, 1973), albeit only slightly. Conversely, responses moving in the other direction should be strengthened.

If, now, we assume that the strengthening of responses in one direction, and/or their weakening in the other, leads to a shift (because of these changes of strength) in the distribution of behavior such that relatively more time is spent in the locally better situation, melioration (and by implication matching) may be viewed as the outcome of the relative strengths of changeover responses within choice situations.

REFERENCES

Baum, W. M. The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 1973, **20**, 137-153.

Baum, W. M. On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 1974, **22**, 231-242.

de Villiers, P. A. Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Fantino, E., & Logan, C. A. *The experimental analysis of behavior*. San Francisco: W. H. Freeman, 1979.

Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1961, **4**, 267-272.

Herrnstein, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 243-266.

Herrnstein, R. J., & Heyman, G. M. Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior*, 1979, **31**, 209-223.

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*. New York: Academic Press, 1980.

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.

Rachlin, H. Contrast and matching. *Psychological Review*, 1973, **80**, 217-234.

Rachlin, H. A molar theory of reinforcement schedules. *Journal of the Experimental Analysis of Behavior*, 1978, **30**, 345-360.

Rachlin, H. Comments on Heyman and Luce. *Animal Learning and Behavior*, 1979, **7**, 267-268.

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.

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

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.

Staddon, J. E. R., & Motheral, S. On matching and maximizing in operant choice experiments. *Psychological Review,* 1978, **85**, 436-444.

Received January 16, 1981
Final acceptance May 5, 1981