MELIORATION AND MAXIMIZATION OF REINFORCEMENT MINUS COSTS OF BEHAVIOR

HARRIE BOELEN

TILBURG UNIVERSITY, THE NETHERLANDS

Eight pigeons were exposed to independent concurrent schedules. Concurrent variable-interval 60-second variable-interval 60-second schedules were presented to one group of four subjects. Following baseline training, a limited hold was added to one of the schedules and the duration of the hold was decreased in successive conditions. Concurrent variable-interval 120-second variable-interval 40-second schedules were presented to another group of four subjects. These subjects were first exposed to decreasing durations of a limited hold in the variable-interval 40-second component. After replication of the baseline, a limited hold in the variable-interval 120-second component was decreased in duration. The initial durations of the holds were determined from the subjects' responding in the baseline conditions. A duration was chosen such that approximately 25% of the scheduled reinforcers would be canceled if responding remained unchanged.

Approximate matching of time proportions and reinforcement proportions was observed when the limited hold was added to the variable-interval 60-second schedule and when the limited hold was added to the variable-interval 40-second schedule. Time proportions were less extreme than reinforcement proportions when the limited hold operated in a variable-interval 120-second schedule. Overall reinforcement rates tended to decrease with continued training in concurrent schedules with a limited hold. Absolute deviations from time matching also decreased. The results provide evidence against the principle of reinforcement maximization, and support Herrnstein and Vaughan's (1980) melioration hypothesis.

Key words: melioration, maximization, concurrent schedules, limited hold, key peck, pigeons

In concurrent schedules of reinforcement, the proportion of time spent responding to an alternative frequently matches the proportion of reinforcers provided by that alternative (Baum, 1979). This has been termed *time matching*. Two hypotheses have been advanced to account for time matching in steady-state performance. Herrnstein and Vaughan's (1980) melioration hypothesis states that time will be reallocated towards alternative i when

$$\frac{R_i}{T_i} - \frac{R_j}{T_j} > 0,$$

where $T_{i,j}$ are the times spent responding

to alternatives, and i, j and $R_{i,j}$ are the rates of reinforcement provided by the alternatives. In addition, it has been shown that approximate time matching in concurrent schedules is consistent with the principle of reinforcement maximization, which states that animals respond in ways that maximize the sum of the reinforcement rates (Houston, 1983).

The reinforcement-maximization principle has been tested in experiments designed to exclude reinforcement maximization as the basis for time matching (Herrnstein & Heyman, 1979; Mazur, 1981; Vaughan, 1981). These studies have produced evidence in favor of the melioration hypothesis and against reinforcement maximization. It could be argued, however, that some function of responding other than overall reinforcement rate was maximized in these experiments. For example, findings

I am indebted to Frans van Haaren for his comments on the manuscript and to Marijke Ponjée and Niek Meulenbroek for assistance in data collection and analysis. Reprints may be obtained from Harrie Boelens, Department of Psychology, Tilburg University, 5000 LE Tilburg, The Netherlands.

obtained in simple and concurrent schedules of reinforcement can be explained by maximization principles if it is assumed that utility consists of both reinforcement and leisure (Green, Kagel, & Battalio, 1982; Rachlin, Battalio, Kagel, & Green, 1981). Such an optimization analysis applies well to the findings of Herrnstein and Heyman (1979), who studied performance in concurrent variable-interval (VI) variable-ratio (VR) schedules. They observed that subjects spent too much time in the VI component, relative to that predicted by reinforcement maximization. Herrnstein and Heyman also noted that local response rates (responses per unit of time spent responding in a component) were higher in the VR than in the VI component. Consequently, increases in the time spent responding in the VI component could be viewed as increased leisure, which may be considered to contribute to a good that is maximized.

The present experiment investigated maximization principles that could possibly govern behavior in concurrent schedules of reinforcement. In addition to reinforcement maximization, two other principles are considered: maximizing of reinforcement minus cost of changing over from one response alternative to the other, and maximizing of reinforcement minus cost of responding. These maximization principles were tested with a procedure that precluded simultaneous time matching and maximization of reinforcement. Concurrent VI VI schedules were used, with a limited hold in one of the two components; effects of decreasing the duration of the limited hold were examined.

The predictions generated by reinforcement maximization and by melioration can be obtained by constructing reinforcement feedback functions, which specify the rate of reinforcement as a function of behavioral variables (Baum, 1973). The derivation of a feedback function for a VI schedule component with limited hold is presented in Appendix A. Figure 1 shows feedback functions corresponding to different limited-hold durations, for a constant probability VI 60-s schedule component considered in isolation.

Fig. 1. Reinforcement feedback functions for a constant probability VI 60-s schedule component shown parametrically for different durations of limited hold. Hold durations are specified in seconds. The local rate of responding in the component, B_1 , was set equal to 1/s, Heyman's inertia parameter (I) was set equal to 3 s, and the duration of a changeover delay to 1 s.

Two parameters in the reinforcement feedback functions correspond to behavioral variables. These are B_1 , which denotes the local rate of responding in the component, and I , Heyman's "inertia" parameter, which scales the subject's tendency to change from one response alternative to the other (Heyman & Luce, 1979). The parameter I is measured in time units; a small value of I corresponds to a high tendency to change. Typical values of B_1 and I have been substituted to obtain the functions shown in Figure 1 ($B_1 = 1/s$; $I = 3s$). For each proportion of time spent in the component, reinforcement rate decreases with decreases in the duration of the hold. Limited holds with a duration greater than the mean interresponse time ($1/B_1 = 1s$, in the present example) almost exclusively cancel reinforcers that were made available while the subject responded on the other response alternative, and exclusive choice of the alternative with the limited hold produces nearly the scheduled rate of reinforcement. Further reductions in the duration of the hold also affect the presentation of reinforcers that have become available while the subject responded on the alternative with the limited hold. Limited holds that are much smaller than the mean interresponse time produce sched-

ules functionally equivalent to extinction.

Figure 2 shows feedback functions for combined outcomes from concurrent VI 60-s VI 60-s schedules with a limited hold in one component but not the other. As the duration of the hold decreases, maximization of overall reinforcement rate occurs first at increasing and then at decreasing values of the proportion of time in the limited-hold component. The change from increasing to decreasing proportions occurs when hold duration is set equal to the mean interresponse time. Melioration predicts decreasing time proportions, because the rate of reinforcement provided by the limited-hold alternative decreases. Similar patterns of effects are predicted for VI schedules with other values.

Fig. 2. Reinforcement feedback functions for combined outcomes from concurrent VI 60-s VI 60-s schedules with a limited hold in the first component. Triangles indicate maxima. Vertical bars indicate the reinforcement rates resulting from time matching. Constants are the same as those used to generate Figure 1.

The concurrent VI VI procedure with a limited hold in one component has an advantage over the concurrent VI VR schedule as a test of melioration versus maximization

principles. Suppose a limited hold is added to one of two concurrent VI components. If the duration of the limited hold is much greater than the mean interresponse time in the component (the reciprocal of the local response rate), then responding in the component containing the limited hold does not make contact with the contingency. Such limited holds cancel only reinforcers for responses that terminate changeovers to the component with the limited hold. Therefore, these limited holds cannot produce differences in local response rates between the alternatives. If the local response rates are equal, then a reallocation of time towards the VI alternative does not increase leisure. As a consequence, the predictions of the melioration hypothesis are different not only from those of reinforcement maximization, but also from those of the maximization of utility, if utility consists of both reinforcement and leisure.

In the present experiment, the effects of introducing a limited-hold contingency in one of two concurrent VI components, and those of subsequently reducing the duration of the hold, were examined. The shortest limited holds to which the subjects were exposed had durations approximately equal to the mean interresponse times. The durations of the limited holds were much greater than the mean interresponse times in most of the other cases. Reinforcement maximization would require, for each of the durations of the limited hold, that the proportion of time in a component be greater when the limited hold was added than when it was not. The concept of melioration predicts that the proportions of time spent in the limited-hold component should decrease with decreasing durations of the hold.

METHOD

Subjects

Eight homing pigeons, between one and five years old, were maintained at 80% of their free-feeding weights with supplemental feedings after each experimental session. Subjects HB33, HB34, HB35, and HB36

were experimentally naive. Subjects HB2, HB3, HB4, and HB7 had prior experience with concurrent schedules of food reinforcement. Subject HB4 died after 20 sessions of training in Condition 10 of the experiment.

Apparatus

Four standard, Lehigh Valley, three-key testing chambers were used. The chambers were 36 cm high, 31 cm long, and 35 cm wide. Only the left and the right keys were operative during this experiment. Both keys were 2.5 cm in diameter and required a force of approximately 0.10 N to be operated. The keys were mounted 9 cm (to the center of the key) from the left and the right walls of the chamber, and they were spaced 16.5 cm apart. The right key could be illuminated by a green light, and the left key by a red light. Access to standard mixed pigeon grain was provided through a 5-cm by 6-cm aperture, centered on the same wall, 11 cm from the floor of the chamber.

A fan provided fresh air and some masking noise. A PDP 8/E minicomputer (Digital Equipment Corporation), located in an adjacent room, and SKED software (Snapper, Stephens, & Lee, 1974) were used to control the experimental procedures and data collection.

Procedure

Subjects HB33, HB34, HB35, and HB36 were first trained to eat from the magazine and to peck the keys. Subsequently, each subject was exposed to a sequence of conditions, in each of which concurrent VI VI schedules were presented. One schedule was assigned to the left key and the other to the right key. The schedules consisted of 15 intervals constructed according to the method of Fleshler and Hoffman (1962) and operated continuously except during reinforcement (3-s presentation of grain). During reinforcement presentation by one schedule, the other schedule continued to operate. A changeover delay (COD) of 1 s specified that reinforcement was not available during 1 s following a response that terminated a changeover from one key to the other.

Table 1 shows the sequences of conditions for subjects HB2, HB3, HB33, and HB34 (Group 1). These subjects were exposed to concurrent VI 60-s VI 60-s schedules in the first condition of the experiment. A limited-hold contingency was added to the right-key schedule in the second condition. The durations of the hold were different for different subjects and were obtained as follows. During the first experimental condition, frequency distributions of set-up times (intervals starting when reinforcement became available and ending with the first subsequent response) were determined for right-key reinforcers. The duration of the hold in the second condition was set equal to the third quartile of this distribution. This procedure guaranteed that approximately 25% of the reinforcers set up by the schedule with limited hold would be canceled if no change in responding occurred. The durations of the hold were gradually reduced in Conditions 3, 4, and 5. Condition 6 replicated Condition 1.

Table 2 shows the sequences of conditions for subjects HB4, HB7, HB35, and HB36 (Group 2). These subjects were exposed to concurrent VI 120-s VI 40-s schedules in the first condition; the VI 40-s schedule was assigned to the right key. This concurrent schedule was selected to provide a baseline with unequal preference between the alternatives. During the first experimental condition, distributions of set-up times in the VI 40-s component were obtained. A limited hold with a duration equal to the third quartile of the distribution was added to the VI 40-s schedule in Condition 2. In Conditions 3, 4, and 5 the duration of the hold was gradually reduced. In Condition 6, which was the same as Condition 1, distributions of set-up times in the VI 120-s component were obtained. A limited hold was then introduced in this component in Condition 7 with a duration equal to the third quartile of the distribution. In subsequent conditions the duration of the hold was reduced until four values of the limited hold had been presented or until exclusive choice had been observed. Thereafter, the concurrent VI

Table 1

Group 1: Sequences of experimental conditions and means from the last five sessions of each condition. Horizontal bars indicate unlimited durations of the hold. (L = left key; R = right key)

Condition	Sessions	VI schedules	Hold durations	Responses		Time	Reinforcers		Changeovers
		(s) L/R	(s) L/R	L	R	(s) R	L	R	to R

Subject HB2									
1	34	60/60	-/-	2966	2270	866	30.2	28.8	278
2	19	60/60	-/5.5	2694	2174	833	29.0	20.6	322
3	15	60/60	-/4.0	2565	2051	722	29.0	14.6	268
4	12	60/60	-/2.5	2874	1815	612	28.0	7.8	267
5	31	60/60	-/1.0	2348	1212	439	28.0	2.8	156
6	10	60/60	-/-	2228	2546	868	28.4	28.6	327
Subject HB3									
1	23	60/60	-/-	977	945	908	27.2	27.0	115
2	19	60/60	-/10.0	1252	1159	715	28.6	22.4	145
3	21	60/60	-/7.5	1942	633	344	28.2	9.4	102
4	13	60/60	-/5.0	2049	417	207	28.4	4.8	70
5	18	60/60	-/2.5	1322	112	57	28.4	1.0	16
6	22	60/60	-/-	1332	1125	847	29.0	27.6	184
Subject HB33									
1	30	60/60	-/-	1493	1490	867	28.2	28.8	215
2	15	60/60	-/5.0	1932	1661	734	29.2	21.0	259
3	19	60/60	-/3.5	1847	1625	692	28.4	16.8	223
4	14	60/60	-/2.0	1839	1106	530	27.6	5.2	173
5	10	60/60	-/0.5	2143	20	11	28.2	0.0	6
6	27	60/60	-/-	2402	1379	573	30.2	27.8	180
Subject HB34									
1	19	60/60	-/-	1235	1472	1032	28.8	28.8	183
2	18	60/60	-/4.5	1256	826	546	27.6	10.2	165
3	20	60/60	-/3.5	1697	160	85	28.8	1.6	38
4	14	60/60	-/2.5	1150	90	61	28.4	1.2	14
5	11	60/60	-/1.5	1461	20	17	28.6	0.0	5
6	33	60/60	-/-	1394	970	794	28.2	28.0	110

120-s VI 40-s baseline was replicated (Condition 9 for Subject HB4, Condition 10 for Subjects HB7 and HB36, and Condition 11 for Subject HB35).

In the final part of the experiment, the subjects in Group 2 were used for a reexamination of the effects of introducing a limited-hold contingency in a component corresponding to a preferred response alternative. After replication of the concurrent VI 120-s VI 40-s baseline, a limited hold of 3 s was introduced in the VI 40-s component (Condition 11 for Subjects HB7 and HB36 and Condition 12 for Subject HB35).

Following stabilization of responding in the fourth replication of the baseline, the schedules were changed to concurrent VI 60-s VI 60-s and a limited hold of 3 s was again introduced in the right-key component (Condition 13 for Subjects HB7 and HB36 and Condition 14 for Subject HB35). This manipulation was arranged to examine the effects of changing schedules and introducing a limited hold on time allocation when reinforcement maximization would specify that time allocation not change. (In the final three conditions to which Subjects HB7, HB35, and HB36 were exposed, the overall

Table 2

Group 2: Sequences of experimental conditions and means from the last five sessions of each condition. Horizontal bars indicate unlimited durations of the hold. (L = left key; R = right key)

Condition	Sessions	VI schedules (s)	Hold durations (s)	Responses		Time (s)	Reinforcers		Changeovers to R
		L/R	L/R	L	R	R	L	R	
Subject HB4									
1	41	120/40	-/-	888	1869	1224	14.8	42.2	114
2	19	120/40	-/5.5	643	2198	1332	13.4	34.0	90
3	16	120/40	-/4.0	718	2611	1306	14.0	34.2	110
4	19	120/40	-/2.5	727	3005	1311	13.8	30.0	105
5	16	120/40	-/1.0	801	2072	1147	13.8	24.0	121
6	12	120/40	-/-	663	2211	1293	14.4	42.4	114
7	11	120/40	12.0/-	468	1773	1402	8.0	43.0	97
8	17	120/40	9.0/-	21	1263	1770	0.0	43.4	5
9	21	120/40	-/-	550	1694	1254	13.8	42.0	98
Subject HB7									
1	33	120/40	-/-	1014	1461	1335	15.0	42.4	174
2	19	120/40	-/3.5	963	1815	1373	14.6	36.0	161
3	16	120/40	-/2.5	961	1781	1352	14.0	32.2	152
4	12	120/40	-/1.5	1226	1732	1114	14.0	20.4	170
5	23	120/40	-/5	1120	1750	919	13.2	14.8	148
6	12	120/40	-/-	1288	2087	1213	14.8	42.6	237
7	12	120/40	6.0/-	1001	2212	1349	8.0	43.4	248
8	13	120/40	4.5/-	811	2163	1444	4.2	42.2	197
9	27	120/40	3.0/-	25	2000	1784	0.0	43.2	12
10	14	120/40	-/-	1194	1866	1315	15.2	43.2	230
11	18	120/40	-/3.0	977	1718	1103	14.2	28.2	162
12	12	120/40	-/-	875	2152	1449	15.0	43.8	167
13	20	60/60	-/3.0	1293	1058	610	27.6	9.8	218
14	14	120/40	-/-	991	2066	1404	15.4	43.8	177
Subject HB35									
1	14	120/40	-/-	616	841	1343	13.6	40.0	107
2	11	120/40	-/5.0	809	1665	1355	14.0	33.8	119
3	11	120/40	-/3.5	791	1360	1255	13.8	27.6	113
4	16	120/40	-/2.5	691	804	996	13.0	20.0	95
5	13	120/40	-/5	995	655	593	13.8	4.0	119
6	21	120/40	-/-	1062	1753	1157	14.4	42.6	171
7	12	120/40	10.0/-	966	1187	1157	10.0	42.4	145
8	12	120/40	7.5/-	675	1439	1377	5.4	42.6	103
9	17	120/40	5.0/-	136	1606	1724	0.6	43.2	17
10	11	120/40	2.5/-	14	1519	1792	0.0	43.4	2
11	15	120/40	-/-	757	1388	1307	13.2	42.2	107
12	11	120/40	-/3.0	662	1235	1272	13.4	29.2	95
13	16	120/40	-/-	910	1130	1151	15.2	43.0	139
14	28	60/60	-/3.0	1174	557	686	27.0	10.4	143
15	23	120/40	-/-	954	1182	1008	14.6	42.4	129
Subject HB36									
1	15	120/40	-/-	1402	3149	1402	14.8	43.4	161
2	10	120/40	-/3.0	1271	2190	1295	14.2	31.8	158
3	14	120/40	-/2.5	1993	2925	1155	14.6	26.2	175
4	14	120/40	-/1.5	1944	2614	956	14.0	21.2	192
5	10	120/40	-/5	1732	2857	914	13.4	14.8	175
6	11	120/40	-/-	2065	3356	1144	15.4	43.4	198
7	13	120/40	8.0/-	1965	3192	1308	12.8	43.2	193
8	11	120/40	6.0/-	1449	2789	1381	7.2	43.4	161
9	25	120/40	4.0/-	8	3183	1796	0.0	43.4	4
10	13	120/40	-/-	1996	3155	1299	15.6	44.0	201
11	10	120/40	-/3.0	2092	3159	1270	15.4	35.4	210
12	10	120/40	-/-	1798	2953	1321	15.4	43.8	208
13	17	60/60	-/3.0	2389	48	13	28.8	0.2	10
14	25	120/40	-/-	1981	2886	1309	15.4	43.4	219

rate of reinforcement would be maximized by spending approximately 80% of the time responding on the right key.)

Conditions were changed when each of the following criteria had been met: (1) proportions of time spent responding showed a range less than or equal to .08 over the last five sessions; (2) at least 10 sessions had been presented; (3) the time proportions, the number of responses on the left key, the number of responses on the right key, and the number of changeovers to the right key did not systematically increase or decrease over the last five sessions.

Experimental sessions ended after 30 min, exclusive of total reinforcement time. Sessions were conducted six days a week.

RESULTS

Tables 1 and 2 summarize the main results for each subject in Groups 1 and 2, respectively. All figures presented, except Figure 5, are based on these data.

Figure 3 shows the proportions of time spent responding on the right key and the proportions of reinforcers produced by right-key responses for each subject in Group 1. Reinforcement and time proportions decreased as the duration of the hold in the right-key component was decreased. Three subjects (HB3, HB33, and HB34) showed approximate matching of time and reinforcement proportions.

Figure 4 shows the same measures for the subjects in Group 2. Introducing the limited hold in the VI 40-s (right-key) component did not consistently affect time allocation; time proportions decreased with further decreases of the duration of the hold. Reduction of hold duration in the VI 120-s component produced increasing time proportions in the VI 40-s component. Approximate time matching was observed when the duration of the hold in the VI 40-s component was reduced. Time proportions were less extreme than reinforcement proportions when a limited hold operated in the VI 120-s component.

Figure 5 shows the shifts in overall rein-

Fig. 3. Proportions of time spent responding on the right key and proportions of reinforcers produced by right-key responses for the subjects trained in concurrent VI 60-s VI 60-s schedules (Group 1). Interconnected data points correspond to a sequence of decreasing durations of the limited hold in the right-key component.

forcement rates and in time proportions that occurred during training in conditions in which a limited hold was in effect. Each point in the figure represents a pair, consisting of a difference between time proportions and a difference between overall reinforcement rates. The difference between time proportions resulted from subtracting the proportion of time spent responding on the limited-hold alternative, averaged over the final three sessions in a condition, from the same measure, averaged over the first three sessions. The difference between overall reinforcement rates resulted from subtracting the sum of the reinforcement rates provided by the left and right keys, averaged over the final three sessions, from

Fig. 4. Proportions of time spent responding on the right key and proportions of reinforcers produced by right-key responses for the subjects trained in concurrent VI 120-s VI 40-s schedules (Group 2). The VI 40-s schedule was assigned to the right key.

the same measure, averaged over the first three sessions.

Increases in overall reinforcement rate occurred in 16 of 50 cases and tended to be smaller than decreases. The largest decreases occurred when concurrent VI 60-s VI 60-s schedules, with a limited hold added to the second schedule, were arranged after performance had stabilized during concurrent VI 120-s VI 40-s schedules (Group 2).

Shifts in absolute deviation from time matching were assessed as follows. First, the number of reinforcers provided by Alternative 1 was divided by the number of rein-

forcers provided by Alternative 2, both numbers being sums over the first three sessions, and the logarithm of the ratio was taken. Second, the time spent in Component 1 was divided by the time spent in Component 2, both times being those spent in the first three sessions. The logarithm of this number was taken and the absolute difference between the two logarithms was obtained. The same computation was made with measures obtained in the last three sessions. Absolute deviations from matching increased in 15 of 43 cases.

Figure 6 shows numbers of changeovers to

Fig. 5. Changes in overall reinforcement rates and changes in proportions of time spent responding on the limited-hold alternative for each subject and each condition with a limited hold. A data point represents the pair of changes that occurred during training in a condition. See text for details.

the right key. For the subjects in Group 1 the number of changeovers decreased after it had been approximately constant in the first two or three conditions. The numbers of changeovers observed in Group 2 were approximately constant when the limited hold was applied to the VI 40-s schedule, but decreased after introduction of the limited hold in the VI 120-s component.

Figures 7 and 8 show local response rates. These rates were calculated by dividing the number of responses on a key by the time spent responding on the key. Figure 7 shows local response rates for the subjects in Group 1. Local response rates in limited-hold (right-key) components tended to increase and then to decrease as the duration of the hold was decreased. Local response rates in the unchanged VI (left-key) component tended to decrease.

Figure 8 shows similar patterns for the subjects in Group 2. Right-key local response rates tended to increase when a limited hold was added to the right-key schedule (VI 40-s), and was decreased in duration. Simultaneously, left-key local response rates tended to decrease. When a

limited hold was added to the left-key schedule (VI 120-s), right-key local response rates decreased, but left-key local response rates did not change systematically.

Local response rates were also computed separately for responding during CODs (COD response rates) and for schedule responding excluding CODs (post-COD response rates). COD response rates varied between 1.7/s and 6.8/s; post-COD response rates varied between 0.4/s and 2.9/s. COD response rates were at least two times greater than post-COD response rates, for each key and each subject in each condition. COD response rates did not show systematic effects with respect to the independent variable. Post-COD response rates showed effects similar to those shown by the local response rates, presented in Figures 7 and 8, for Subjects HB3, HB4, and HB35. This measure did not show systematic effects for the other subjects.

DISCUSSION

The effects of reducing the duration of reinforcement availability in one of two concurrent VI schedule components were examined. Both reinforcement proportions and time proportions decreased as the limited-hold duration decreased. Approximate matching of time and reinforcement proportions was observed. Therefore, the present results provide evidence in favor of the melioration hypothesis, which states that time is reallocated towards alternative i when $(R_i/T_i) - (R_j/T_j) > 0$, which leads to time matching in the steady state.

The finding that time proportions decreased as the hold was decreased in duration provides evidence against the principle of reinforcement maximization, which states that animals behave in ways that maximize the overall rate of reinforcement. Reinforcement maximization predicted, for each of the limited holds, that the proportion of time in a schedule component should be greater when the limited hold was added than when it was not.

Overall reinforcement rates produced by

Fig. 6. Numbers of changeovers to the right key for each subject in Group 1 (left panel) and for each subject in Group 2 (right panel).

Fig. 7. Local response rates (number of responses on a key divided by the time spent responding on the key) for each subject in Group 1. No data point is shown when the subject spent less than 20 s per session responding on the key.

Fig. 8. Local response rates for each subject in Group 2.

initial performance in a condition tended to be greater than overall reinforcement rates produced by steady-state responding (Figure 5). Increases occurred in 16 of 50 cases and were smaller than decreases. Because subjects experienced reinforcement rates that were greater than those produced by their steady-state behavior, it is implausible that they used an optimum seeking ("hill-climbing") technique.

It will be shown in what follows that two other maximization principles also do not provide satisfactory explanations of the present findings. First, consider the maximization of $f(R) - g(S)$, where f and g are nondecreasing functions of R , the overall rate of reinforcement, and S is the rate of changing over between response alternatives. The function $f(R)$ could represent the profits produced by reinforcement and the function $g(S)$ the costs of changing over between alternatives. The relation between S and the proportion of time spent in the limited-hold component, p , is given by $S = 2p(1 - p)/I$, where I represents the tendency to changeover between alternatives. S is maximal for $p = 0.5$ and decreases to 0.0 as exclusive choice of either component is approached. In the present experiment, changeover rates were approximately constant for Group 2, when a limited hold operated in the VI 40-s component, and in initial conditions for Group 1. Simultaneously, time proportions decreased. Compare this performance with that of a reinforcement maximizer that shows a constant tendency to changeover (constant I). The reinforcement maximizer will spend more time in a schedule component when a limited hold is added to the schedule. Its changeover rate S and the corresponding cost $g(S)$ will decrease because p deviates from 0.5. Therefore, if $f(R) - g(S)$ is the good that is maximized, then the subjects could have done better by following a reinforcement maximizing strategy, when a limited hold was added to the right-key schedule. This would have produced both increases of $f(R)$ and decreases of $g(S)$.

Second, consider the maximization of

$f(R) - h(B)$ where h is a nondecreasing function of B , the overall rate of response. The function $h(B)$ could represent the costs of responding, or the function $-h(B)$ could represent the value of leisure. One function of the type $f(R) - h(B)$ is $x_1 E^y + x_2 L^y$, which Rachlin et al. (1981) successfully applied to the findings obtained by Herrnstein and Heyman (1979). Here x_1 , x_2 , and y are positive constants, E represents total reinforcement time, and L represents total leisure. (The function $x_1 E^y + x_2 L^y$ can be shown to be of the type $f(R) - h(B)$ by setting $f(R) = x_1 R^y$, $h(B) = -x_2(1 - dB)^y$, where d is the duration of a response and the time unit equals the duration of an experimental session.)

In the present study the local response rate in the limited-hold component tended to increase with decreasing hold duration. Simultaneously, the local response rate on the other alternative decreased. COD response rates were at least twice as great as post-COD response rates; these measures did not vary systematically when the duration of the hold was decreased. This implies that the effects shown by the local response rate were produced by changing durations of stay (interchangeover times) at the response alternatives.

If COD response rates and post-COD response rates are invariant and if COD response rates are greater than post-COD response rates, then overall response rates are decreased by decreasing changeover rates. Therefore, if $f(R) - h(B)$ is the good that is maximized, then a reinforcement maximizer does better than the present subjects did when a limited hold was added to the right-key schedule. A reinforcement maximizer would earn more reinforcers and would reduce costs of responding by decreasing rates of changeover.

The melioration and maximization predictions are not dependent on the assumptions used to construct the particular reinforcement feedback functions shown in Figures 1 and 2. One of the assumptions (see Appendix A) stated that the interresponse times and interchangeover times have a fixed

duration in the steady state. Predictions similar to those presented in Figure 2 result if it is assumed that (1) responding in the schedule components is governed by Poisson processes, and that (2) subjects always outstay a COD, but then change to the other response alternative according to a Poisson process. In addition, it is not essential that both schedules halt when reinforcement is presented. In the present experiment, one schedule continued to operate when the other presented reinforcement. A correction of the feedback functions that takes the extra schedule operation time into account does not lead to changes in the predictions. Reinforcement maximization predicts, on the basis of each of the reinforcement feedback models, for each of the limited holds used in the present experiment, that the proportion of time spent responding on a schedule should be greater when the limited hold is added to the schedule than when it is not.

In conclusion, the present results provide evidence for Herrnstein and Vaughan's (1980) melioration hypothesis, and suggest that time allocation in concurrent schedules is not governed by maximization of reinforcement minus costs of behavior.

REFERENCES

- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, *20*, 137-153.
- Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, *32*, 269-281.
- Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, *5*, 529-530.
- Green, L., Kagel, J. H., & Battalio, R. C. (1982). Ratio schedules of reinforcement and their relation to economic theories of labor supply. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts*. Cambridge, MA: Ballinger.
- Herrnstein, R. J., & Heyman, G. M. (1979). Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior*, *31*, 209-223.
- Herrnstein, R. J., & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143-176). New York: Academic Press.
- Heyman, G. M., & Luce, R. D. (1979). Operant matching is not a logical consequence of maximizing reinforcement rate. *Animal Learning and Behavior*, *7*, 133-140.
- Houston, A. I. (1983). Optimality theory and matching. *Behaviour Analysis Letters*, *3*, 1-15.
- Mazur, J. E. (1981). Optimization theory fails to predict performance of pigeons in a two-response situation. *Science*, *214*, 823-825.
- Parzen, E. (1962). *Stochastic processes*. San Francisco: Holden-Day.
- Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, *4*, 371-417. (Includes commentary)
- Snapper, A. G., Stephens, K. R., & Lee, D. M. (1974). *The SKED software system*. Kalamazoo, MI: State Systems, 1974.
- Vaughan, W., Jr. (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior*, *36*, 141-149.

Received November 22, 1983
Final acceptance May 3, 1984

APPENDIX A

Consider the reinforcement feedback provided by a constant-probability VI schedule, with a scheduled reinforcement rate r and a limited hold of duration H . Let this be Schedule 1. Suppose that one other schedule operates simultaneously and that each changeover is followed by a changeover delay of duration D . A reinforcement feedback function for Schedule 1 can be derived from the following assumptions. First, the probability that a response is reinforced after a reinforcer has been canceled during the immediately preceding interresponse time or interchangeover time is negligible. Second, the interresponse time $1/B_1$ and the interchangeover times $T_{1,2}$ have a fixed duration in the steady state. Third, the first response following a changeover delay is made immediately after the changeover delay expires. Finally, the duration of a changeover to the other alternative equals the interresponse time.

The first assumption leads to the following function $q(t)$, which specifies the probability of reinforcement as a function of time since the last response:

$$q(t) = 1 - e^{-rt}, 0 \leq t \leq H; \\ = e^{-rt}(e^{rH} - 1), t > H. \quad (\text{A.1})$$

Therefore, the probability q_1 that the first response following a COD period is reinforced is given by:

$$q_1 = 1 - e^{-r\{(1/B_1) + T_2 + D\}}, (1/B_1) + T_2 + D \leq H; \\ = e^{-r\{(1/B_1) + T_2 + D\}}(e^{rH} - 1), (1/B_1) + T_2 + D > H. \quad (\text{A.2})$$

The probability of reinforcement for a subsequent response q_2 , is given by:

$$q_2 = 1 - e^{-r/B_1}, 1/B_1 \leq H; \\ = e^{-r/B_1}(e^{rH} - 1), 1/B_1 > H. \quad (\text{A.3})$$

An interchangeover time in schedule Component 1 consists of a changeover delay, followed by $B_1(T_1 - D)$ time segments, each with duration $1/B_1$. The first $B_1(T_1 - D) - 1$ of these segments are terminated by a response to Alternative 1. Each of these responses is reinforced with probability q_2 . (The last segment is the time taken by a changeover to Alternative 2.)

The asymptotic reinforcement rate R_1 is then given by:

$$R_1 = \frac{q_1 + q_2\{B_1(T_1 - D) - 1\}}{T_1 + T_2}. \quad (\text{A.4})$$

The reinforcement rate R_1 can be written as a function of p , the proportion of time spent responding on Schedule 1, by substituting $T_1 = \frac{I}{1-p}$, and $T_2 = \frac{I}{p}$, where I represents Heyman's "inertia" parameter (Heyman & Luce, 1979).

It can be shown that the asymptotic probability of reinforcement as a function of time since the last response equals $H/[H + (1/r)]$ (Parzen, 1962). Therefore, one can improve equation (A.4) by using:

$$q(t) = 1 - e^{-rt}, 0 \leq t \leq H; \\ = e^{-rt}(e^{rH} - 1), H < t \leq T; \\ = \frac{H}{H + (1/r)}, t > T, \quad (\text{A.5})$$

where T , which is equal to $(1/r)\ln(1 + 1/rH)$ ($e^{rH} - 1$), is the time when $q(t)$ as defined by (A.1) intersects the horizontal line $H/[H + (1/r)]$ for the second time. The feedback function $R_1(p)$ that results when the probabilities q_1 and q_2 are determined from (A.5) is slightly different from (A.4) for small values of p , but does not lead to other matching and maximizing predictions.