JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1995, 64, 277–297 NUMBER 3 (NOVEMBER)

HOW TO TEACH A PIGEON TO MAXIMIZE OVERALL REINFORCEMENT RATE

GENE M. HEYMAN AND LAWRENCE TANZ

HARVARD UNIVERSITY

In two experiments deviations from matching earned higher overall reinforcement rates than did matching. In Experiment 1 response proportions were calculated over a 360-response moving average, updated with each response. Response proportions that differed from the nominal reinforcement proportions, by a criterion that was gradually increased, were eligible for reinforcement. Response proportions that did not differ from matching were not eligible for reinforcement. When the deviation requirement was relatively small, the contingency proved to be effective. However, there was a limit as to how far response proportions could be pushed from matching. Consequently, when the deviation requirement was large, overall reinforcement rate decreased and pecking was eventually extinguished. In Experiment 2 a discriminative stimulus was added to the procedure. The houselight was correlated with the relationship between response proportions and the nominal (programmed) reinforcement proportions. When the difference between response and reinforcement proportions met the deviation requirement, the light was white and responses were eligible for reinforcement. When the difference between response and reinforcement proportions failed to exceed the deviation requirement, the light was blue and responses were not eligible for reinforcement. With the addition of the light, it proved to be possible to shape deviations from matching without any apparent limit. Thus, in Experiment 2 overall reinforcement rate predicted choice proportions and relative reinforcement rate did not. In contrast, in previous experiments on the relationship between matching and overall reinforcement maximization, relative reinforcement rate was usually the better predictor of responding. The results show that whether overall or relative reinforcement rate better predicts choice proportions may in part be determined by stimulus conditions.

Key words: matching law, rational choice, maximizing, overall reinforcement rate, relative reinforcement rate, choice, concurrent variable-interval schedule, key peck, pigeons

Imagine the following choice experiment. Responses at one alternative are reinforced according to the passage of time (variable-interval schedule); responses at the other alternative are reinforced probabilistically (variable-ratio schedule). For one group of subjects, the contingencies are represented schematically, say as a graph that displays the reinforcement rates as a function of response and changeover rates. The subjects are then asked, on the basis of the graph, to choose one of the many possible reinforcement-rate combinations, for example, 30 ratio reinforcers and 20 interval reinforcers versus 40 ratio reinforcers and 18 interval reinforcers. For another group of subjects, there are two physically separate alternatives, and at each, reinforcers occur intermittently. That is, this group is in a typical experiment. Do subjects in these two experiments select the same combination of ratio and interval reinforcement rates, and if not, how would their choices differ?

Analogues of the first experiment can be found in any economics textbook. A choice between two options is represented as a choice between different combinations of each option. When this framework is applied to concurrent reinforcement schedules, the implication is that the subject will choose the reinforcement-rate combination that was largest, all else being equal. Thus, in the above example, 58 dominates 50.

Various versions of the second experiment have been conducted with humans, rats, and pigeons (e.g., De Carlo, 1985; Green, Rachlin, & Hanson, 1983; Herrnstein & Heyman, 1979; Herrnstein, Loewenstein, Prelec, & Vaughan, 1993; Heyman & Herrnstein, 1986; Hinson & Staddon, 1983; Mazur, 1981; Savastano & Fantino, 1994; Williams, 1985). In these studies, the subjects typically did not respond so as to maximize the nominal overall reinforcement rate. Instead, behavior usually stabilized when response proportions approximated reinforcement proportions (but see

---

The preparation of this manuscript was supported by NSF Grant 31-702-7510-2-30 to the first author.

Correspondence may be addressed to Gene Heyman, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138 (E-mail:gmh@wjh12.harvard.edu).

Green et al., 1983). This is the well-known matching law (Herrnstein, 1961, 1970). It has a number of mathematical forms, but the simplest is adequate for this report:

$$B_1/(B_1 + B_2) = R_1/(R_1 + R_2), \qquad (1)$$

where B and R refer to responses and reinforcers, respectively, under different alternatives.

There are several points to be made about the two versions of the concurrent interval-ratio experiment. First, for any reinforcement schedule there are several (if not a large number) of possible reinforcement contingencies. For example, in the concurrent variable-interval (VI) variable-ratio (VR) "thought problem" that introduced this paper, behavior may come under the control of overall reinforcement rates, local reinforcement rates, or moment-to-moment reinforcement probabilities. This implies that the same nominal setting (e.g., concurrent VI VR) may yield different behavioral outcomes (as a function of the contingency or blend of contingencies that proved to be most effective). Second, the thought problem suggests that the manner in which a reinforcement schedule is presented will determine which contingency proves to be the effective one. Third, the two preceding points imply that whether or not economic rationality predicts choice will in part depend on contextual features, such as stimulus conditions.

An experiment by Savastano and Fantino (1994), in which the subjects were college students, supports this last point and provides an example of what we mean by "context." In some conditions of a concurrent interval-ratio procedure, the students were given timers that signaled how long a schedule had been running. While the timers were on, choice proportions for 1 of the 6 subjects shifted significantly in the direction of the maximizing predictions. Although the effect was not robust, it does suggest that stimulus conditions may be important.

The influence of context on choice is not restricted to humans. Pigeons shifted from matching toward maximization in experiments in which changes in choice proportions had a relatively immediate effect on overall reinforcement rate (Davison & Alsop, 1991; Silberberg & Ziriax, 1985). In these experiments, reinforcement was dependent on choice proportions as calculated over a "time window" (e.g., a 2-min sample). The basic finding was that the duration of the time window influenced whether relative or overall reinforcement rate better predicted response proportions. For longer time windows, relative reinforcement rate was the better predictor, whereas for shorter time windows, overall reinforcement rate was the better predictor. An important implication of this finding is that the mechanisms that mediated adaptation to the schedules were compatible with both matching and maximizing, with the outcome depending on the temporal domain over which the contingency operated.

In the two experiments described in this report, we manipulated the relationship between overall reinforcement rate and choice in concurrent interval schedules, using a procedure somewhat similar to that of Davison and Alsop and an earlier study by Heyman (1977).[1] The results indicate that discriminative stimuli can determine whether relative or overall reinforcement rate better predicts choice.

EXPERIMENT 1

In Experiment 1 deviations from matching earned higher overall rates of reward than did matching. This was arranged by keeping track of response proportions as calculated over a moving average of the last 360 responses. When response proportions deviated from reinforcement proportions, responses were eligible for reinforcement. When response proportions approximated reinforcement proportions, responses were not eligible for reinforcement. The moving window was set at 360 pecks because pilot work showed that with this size sample it was possible to shape deviations from matching (Heyman, 1977). The window was updated with each peck, as described in more detail below.

METHOD

Subjects

Four male White Carneau pigeons at approximately 80% of their free-feeding weights served as subjects. The birds had been sub-

---

[1] Heyman, G. M. (1977). *Reinforcing deviations from matching.* Paper presented at the meeting of the Eastern Psychological Association, Boston.

jects in previous operant conditioning experiments, and 3 had served in a pilot study similar to the one reported here (Pigeons 59, 62, and 489).

Apparatus

A standard two-key experimental chamber (31 cm high, 33 cm long, and 29.5 cm wide) was used. On the front wall were two response keys (19 mm diameter) and an opening that provided access to the grain hopper. The keys were set 14.5 cm apart and at a height of about 22 cm from the floor. During experimental sessions, the response keys were illuminated from behind with white light and were operated by a force of 0.15 N or more. Effective responses produced auditory and visual feedback (a brief relay click and light flicker). The opening of the grain hopper was 8.9 cm from the floor and midway between the response keys. On the ceiling were two small houselights (28 V DC). The experimental chamber was enclosed in a sound-attenuating box. A fan and white noise generator masked extraneous sounds. Data collection and the presentation of experimental events were controlled by a Digital Equipment Corporation® computer.

Procedure

Baseline conditions. At each key, responses were reinforced at varying intervals. However, in order to insure that the overall relative reinforcement rate remained approximately constant, a single-timer procedure was used (Stubbs & Pliskoff, 1969). The average duration was 25 s, the shortest interval was about 1 s, and the longest was about 75 s. When an interval elapsed, a probability device selected which key would provide the next reinforcer (4-s access to the grain hopper). The next response at the selected key operated the grain hopper and restarted the timer. Thus, each reinforcer was delivered in the order it was assigned, insuring that obtained and programmed reinforcement proportions would be the same. The probability ratios were 1:1, 3:1, and 9:1 (resulting in nominal average interreinforcement intervals of 50 s and 50 s, 33.3 s and 100 s, and 27.8 s and 250 s).

A switch from one to the other schedule (changeover response) initiated a 4-s changeover delay (COD) during which reinforcers were not delivered. This contingency is said to prevent the adventitious strengthening of switching (Findley, 1958; Herrnstein, 1961). Baseline conditions remained in effect for 30 sessions.

Experimental conditions. In experimental sessions there was the added contingency that deviations from matching earned higher overall reinforcement rates. This sort of contingency requires a population of responses for calculating response proportions and a criterion for determining what counts as a deviation from matching. The population was the just previous 360 responses. For instance, at the 720th response of the session, response proportions were calculated on the basis of cardinal Responses 321 to 720, and at the 721st response of the session, response proportions were calculated on the basis of Responses 322 to 721, and so on. (The session began with a sample of 360 responses based on the last 360 responses of the previous session.)

The deviation requirement went into effect once a timer had elapsed and set up a reinforcer. To see how this worked, consider the contingencies for the 2 subjects under the VI 33.3-s VI 100-s schedule in the first experimental condition. Perfect matching was 75% or 270 left pecks (recall that the sample size was 360 pecks) and the deviation requirement was ±5% (±18 pecks). Assume that a reinforcer had just been assigned to the left key. The next response to this key would be reinforced if the number of left responses in the just previous 360 responses was less than 252 or greater than 288 (and the COD had timed out). However, if the choice proportion was within the penalty zone (left counts from 252 to 288), the response to the primed key was not reinforced and the timer was restarted with a new interval. That is, matching "flushed out" the reinforcer, thereby reducing reinforcement rate. (The question of whether reinforcement proportions remained constant is addressed in the Results section.)

For all subjects the deviation requirement began at ±5% and was increased in 5% (18-response) increments. For example, in the second condition for the 2 birds at the 3:1 reinforcement ratio, responses to a primed key were reinforced when left counts were less than 234 or greater than 288. Each deviation requirement was kept in effect for at least 15 sessions and until relative response

Table 1

Experiment 1: Order of conditions, number of sessions, and obtained relative reinforcement rate.

Pigeon	Condition (penalty zone)	Sessions	Obtained left reinforcement (%)
45 (VI 50 s VI 50 s)	0 (baseline)	30	49
	10%	26	48
	15%	23	46
	20%	52	53
	25%	15	53
	30%	39	51
	35%	25	52
62 (VI 250 s VI 27.8 s)	0 (baseline)	30	9
	10%	29	9
	15%	43	13
	20%	37	9
	25%	38	11
	30%	22	12
	35%	34	10
	40%	15	9
	45%	20	11
59 (VI 33.3 s VI 100 s)	0 (baseline)	30	76
	10%	37	71
	15%	40	75
	20%	17	74
	25%	35	76
	30%	22	77
	35%	—	—
489 (VI 33.3 s VI 100 s)	0 (baseline)	30	75
	10%	28	73
	15%	30	72
	20%	48	75
	25%	22	72
	30%	19	75
	35%	15	81
	40%	—	—
	45%	—	—

rates seemed to be stable. (See Table 1 for the number of sessions in each condition.) Sessions ended after 60 primed reinforcers or 40 min, whichever came first.

RESULTS

Figure 1 shows the effect of reinforcing deviations from matching on response and time allocation. The data points are averages, calculated over the last 10 sessions of a condition (except for the final condition, as described below), and the hatched areas indicate response-rate proportions, as calculated over the 360-response window, that were not eligible for reinforcement (penalty zones). (See Appendix A for time and response counts and changeovers.) The obtained reinforcement proportions are listed in Table 1. They did not shift markedly from the programmed values during the course of the study.

Figure 1 shows that in the first four experimental conditions, the larger the deviation requirement, the greater the difference between response and reinforcement proportions. For example, the average differences between response and obtained reinforcement proportions were 3% in baseline (no penalty zone) and then 5%, 9%, 13%, and 16% as the deviation requirement was increased from 10% to 25%. However, the effectiveness of the contingency was limited. Once the penalty zone reached 25%, it was not possible to push response proportions further from matching. Consequently, with further increases in the deviation requirement, responses were no longer eligible for reinforcement, and responding was extinguished. (In each subject's last condition, the measures were taken from the first 10 sessions, because responding had been extinguished.)

Figure 1 (filled circles) also shows the overall proportions of time spent responding at the left key. In previous studies, time allocation and response allocation approximated one another (e.g., Baum & Rachlin, 1969; Stubbs & Pliskoff, 1969; Williams, 1988). However, for the 3 subjects with unequal left and right reinforcement schedules (Pigeons 62, 59, and 489), time allocation and response allocation diverged during experimental conditions. Response proportions shifted from matching, as just described, whereas time proportions typically remained within 5% of matching. However, for the subject at equal left and right schedules (Pigeon 45), this pattern did not hold. For this subject, deviations from matching as measured by time allocation were greater.

Local response rate is defined as the rate of responding while at an alternative (e.g., responses on the left key divided by time spent at the left key). Figure 2 shows this measure as a function of the penalty zone (data averaged from the last 10 sessions, as in Figure 1). Comparing across birds, local rates at the two keys did not systematically differ in baseline sessions. However, for the pigeons at un-

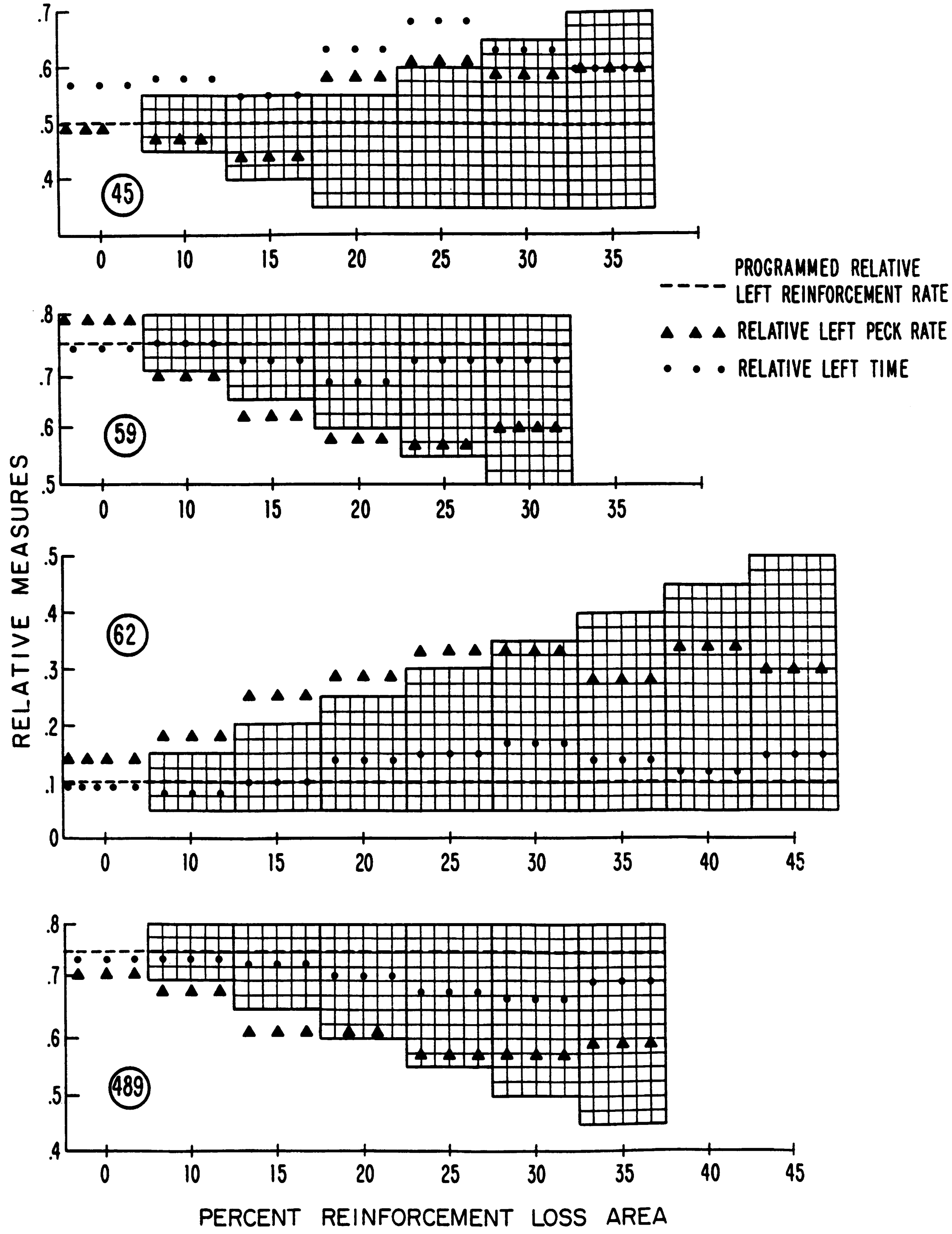

Fig. 1. The *x* axis shows the deviation from matching requirement, expressed as a percentage of the 360-response sample used in the experimental contingency. The *y* axis shows overall response, time, and reinforcement proportions. The hatch lines show the choice proportions, as calculated over the 360-response moving average, that were not eligible for reinforcement. The data points were averaged from the last 10 sessions of each condition, except for the last condition (see text).

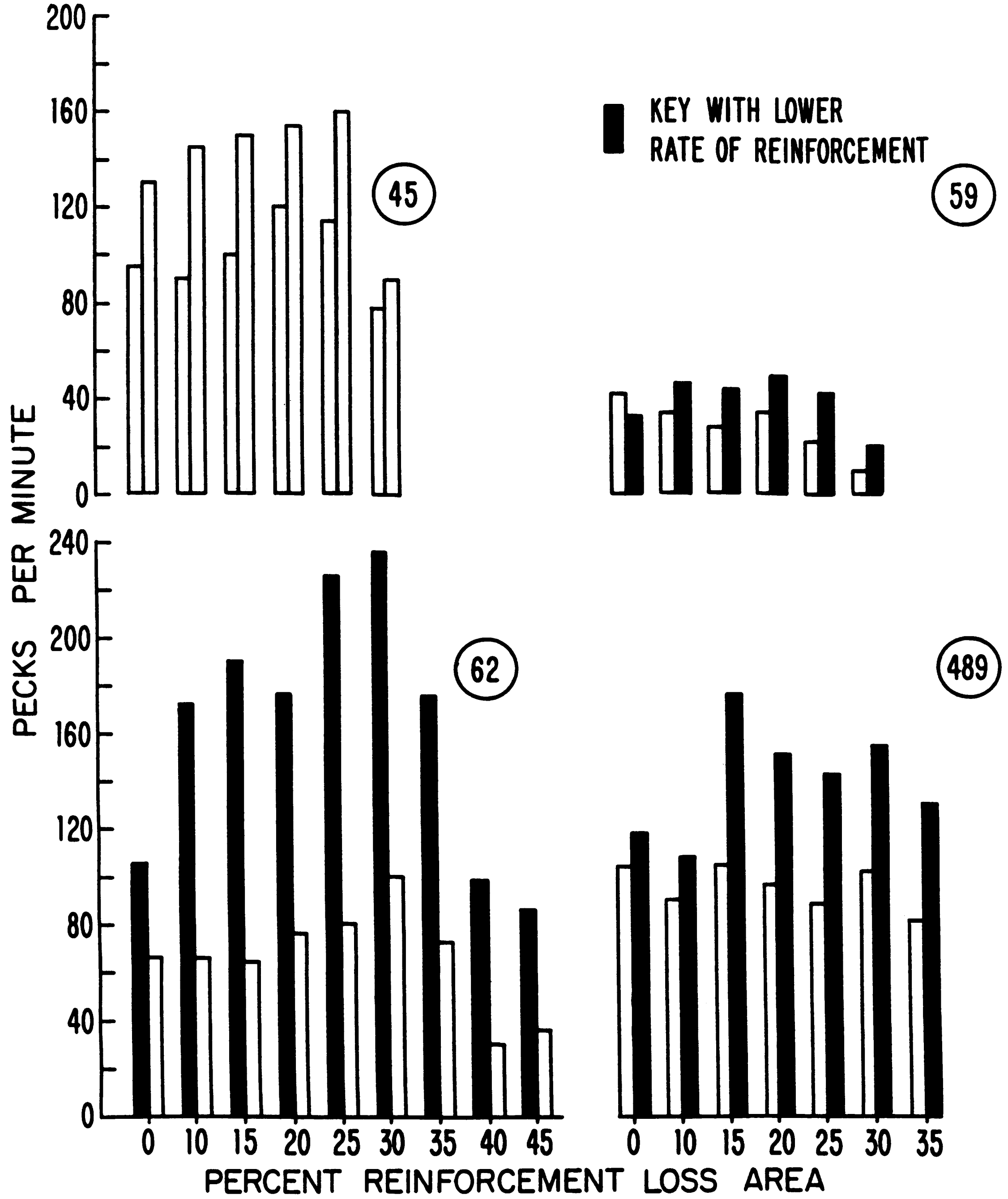

Fig. 2. Local response rates as a function of the deviation requirement. Pigeon 45 was on a VI 50-s VI 50-s schedule, Pigeons 59 and 489 were on a VI 33.3-s VI 100-s schedule, and Pigeon 62 was on a VI 250-s VI 27.8-s schedule.

equal left and right schedules (59, 62, and 489), local response rates at the key with the lower reinforcement rate increased during the first several experimental conditions. However, in the last one or two conditions, when pecking began to extinguish, overall and local rates declined relative to baseline. For the subject at equal left and right schedules (45), the pattern was different. Local response rates increased somewhat at both keys.

Figure 3 shows the relationship between deviations from matching and overall reinforcement rate for three conditions: the first, the fourth (in which each subject stabilized at or within 1% of its maximum deviation), and the last (see Table 1 for response-count boundaries of the penalty zone). Every session from the three conditions was used, and sessions were grouped by deviations from matching, using 2% bins. The *y* axis shows the median number of reinforcers for each bin.

In general, the greater the deviation, the greater the reinforcement rate (as planned). For example, when the deviation requirement was 10% (first condition), a 5% average deviation from matching earned about 20 reinforcers per session and a 10% average deviation earned about 40 reinforcers per session. However, comparison of Figures 3 and 1 shows that the birds' performances did not stabilize at choice proportions that produced the highest overall number of reinforcers. For example, in the first condition, Pigeon 45 stabilized at a deviation (3%) that produced 41 reinforcers per session, even though there had been sessions in which it earned as many as 80 reinforcers (for average deviations of 18%).

If choice proportions, as calculated over the most recent 360 responses, failed to exceed the deviation requirement, then reinforcers were not delivered, and there could be no correlation between choice and overall reinforcement rate. Figure 3 shows that this is what happened in each subject's last condition. For example, the minimum deviation requirement in a last condition was 30%, and Figure 3 shows that there were no sessions in which the average overall choice proportions differed by more than 30%.

Although we have described the contingency in terms of the relationship between response and reinforcement proportions, increases in run length (number of responses between changeovers) could in principle affect whether or not a reinforcer was delivered. For example, if run lengths were at least 360 responses long, response proportions, when calculated in terms of the moving window, would vary widely and often exceed the penalty zone, regardless of the overall session response proportion. Thus, it is possible that subjects would learn to change over less frequently. However, Figure 4 shows that, except for Pigeon 45, run lengths typically decreased when a penalty zone was in effect. For Pigeon 45, the subject on equal left and right schedules, run lengths increased.

DISCUSSION

The contingency used in this experiment more effectively shaped deviations from matching than did similar studies that used relatively large averaging windows (e.g., Davison & Kerr, 1989; Vaughan, 1981). The difference in results may be a function of the nature of the averaging window used to shape response proportions. Previous studies used temporally defined averaging windows, whereas we used a response-defined window. However, even with a response-defined window, it was not possible to shape any arbitrarily selected response proportion. This constraint may reflect limitations in the procedure or, alternatively, limitations in what the subjects could learn. In favor of the procedural account are experiments that more successfully shaped deviations from matching. Davison and Alsop (1991) and Silberberg and Ziriax (1985) shaped response proportions to arbitrarily selected values when the window size was 6 s or less. Similarly, the contingency may have been more effective if overall reinforcement density were greater. However, Figure 3 shows that the reinforcement gradient was not unusually shallow. Even in the condition in which pecking was extinguished, there were sessions in which the pigeons earned reinforcement rates that had reliably maintained responding in other experiments (e.g., Findley, 1958).

For the 3 subjects at unequal left and right schedules, time allocation was not as affected by the contingency as was response allocation. This may reflect the fact that the contingency did not specify how time was to be spent, or, alternatively, time allocation may have been more resistant to the shaping procedure. This could be tested by arranging a procedure that differentially reinforced time allocation independently of response allocation. Experiment 2 follows up on the idea that the limits on deviations from matching could have been overcome if the correlation

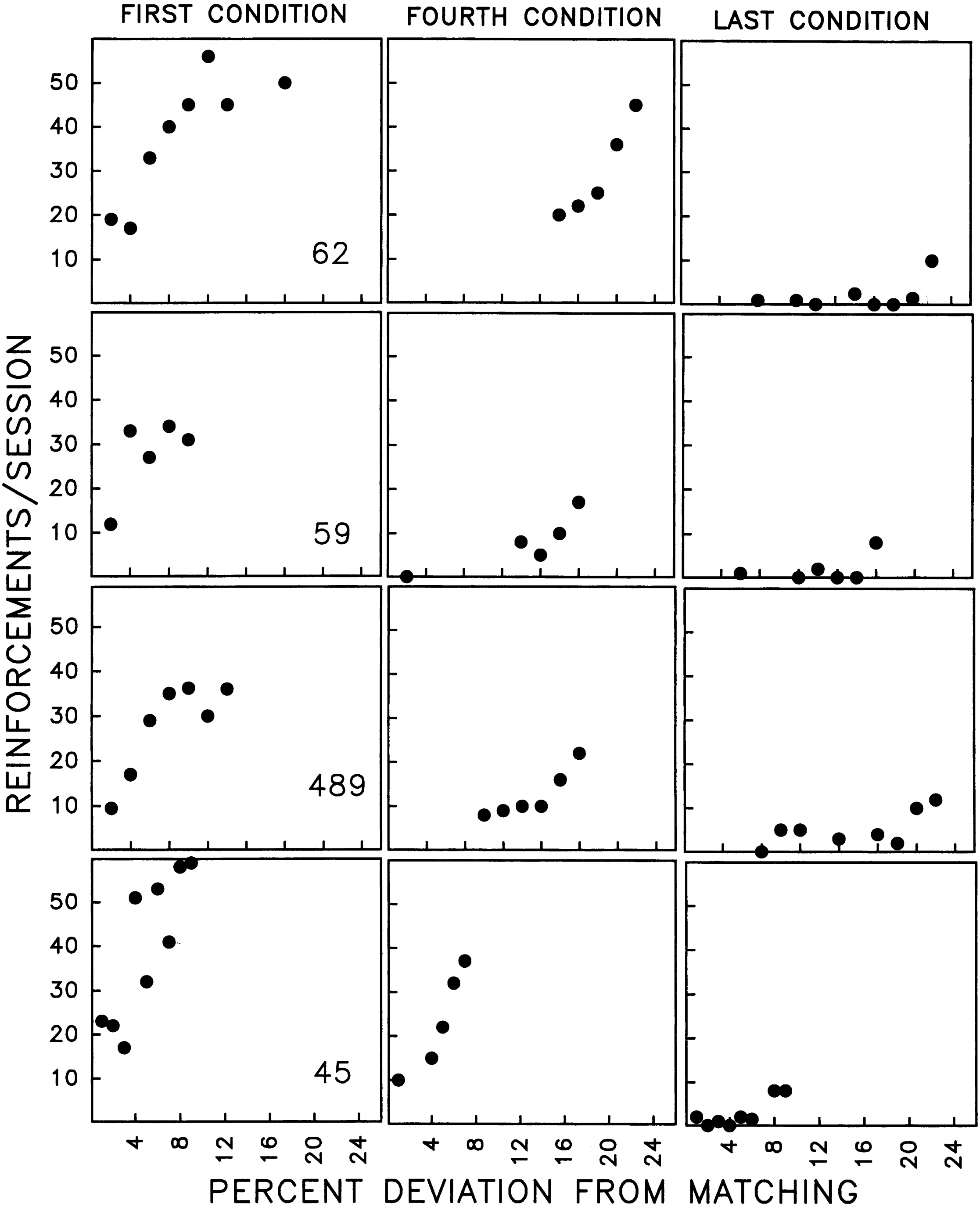

Fig. 3. The number of reinforcements per session as a function of the degree to which choice proportions differed from the nominal reinforcement proportions. Each point gives the median number of reinforcers for the bin indicated by the *x* axis. All sessions for each of the first, fourth, and final experimental conditions are included.

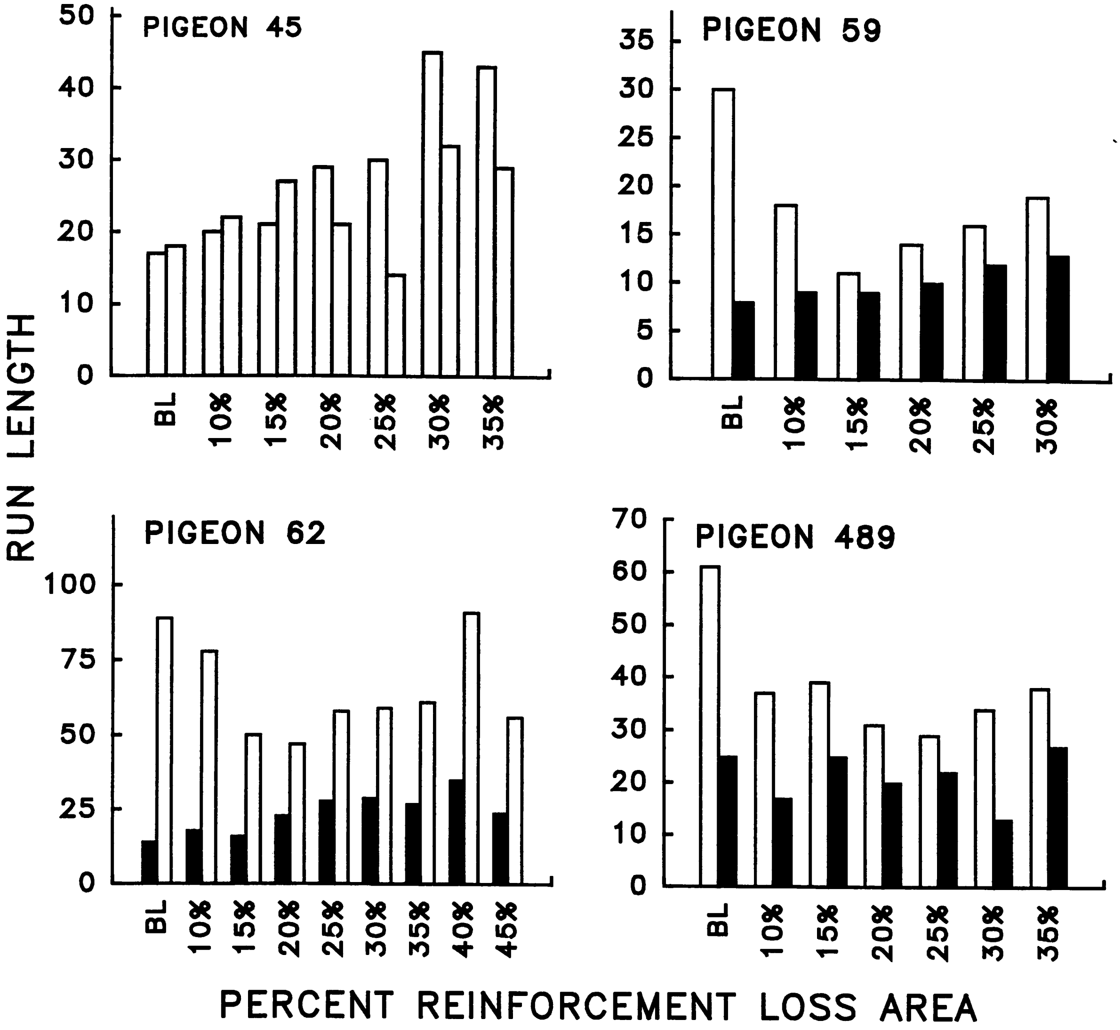

Fig. 4. Number of responses between changeovers (run lengths) as a function of the deviation requirement. The data are averaged from the last 10 sessions of a condition, except for the last condition (see text). Open bars represent the key with the higher rate of reinforcement; solid bars represent the key with the lower rate of reinforcement.

between responding and overall reinforcement rate had been stronger.

EXPERIMENT 2

In Experiment 1, as in previous studies, changes in overall reinforcement rate were not signaled by external stimuli. In contrast, the component reinforcement rates were associated with clear landmarks (e.g., location). This should not be surprising because every study in which overall and relative reinforcement rate predicted different outcomes, stimulus conditions were correlated with relative reinforcement rates (or, equivalently, local reinforcement rates). Experiment 2 was designed differently. The houselight changed color as a function of changes in overall reinforcement rate. To our knowledge, this is the first study in which discriminative stimuli explicitly signaled changes in overall reinforcement rate.

METHOD

Subjects

Three male homing pigeons were maintained at approximately 85% of their free-feeding weights. The pigeons had been subjects in previous reinforcement-schedule experiments.

Apparatus

A standard two-key operant conditioning chamber was used. The right response key was illuminated from behind by a red light, and the left response key was illuminated in the same manner by a green light. A force of about 0.15 N or greater operated the response keys. On the ceiling of the experimental chamber were a blue light and a white light. During experimental sessions, one or the other was on, depending on the subject's behavior (see below). Other features of the apparatus were as described in Experiment 1.

Procedure

Baseline. Responses on each key were occasionally reinforced according to a VI schedule. As in Experiment 1, the schedules were arranged so that relative reinforcement frequency remained approximately constant (Stubbs & Pliskoff, 1969). However, in this experiment, a somewhat different method was used. There were two timers. When an interval timed out on one key, thereby setting up a reinforcer, the timer at the other key stopped. Thus, each reinforcer was delivered in the order set by the schedules, insuring that obtained and programmed reinforcement proportions were the same (all else being equal, but see below). The nominal average interreinforcement intervals were 30 s at each key for Pigeon 1 (1:1), 60 s and 20 s for Pigeon 2 (1:3), and 150 s and 16.6 s for Pigeon 3 (1:9). These mean values were selected so that programmed overall reinforcement rates were four per minute. The list of intervals for each schedule approximated a Poisson distribution, with the longest interval set at about four times the mean (Fleshler & Hoffman, 1962). The changeover delay was set at 1.5 s, and reinforcement consisted of 3-s access to the grain hopper. Session length was 30 min. These conditions were kept in effect for 50 sessions.

Experimental conditions. As in Experiment 1, deviations from matching earned higher rates of reinforcement. However, there were several important modifications of the procedure. First, the sample size for calculating deviations from matching was varied. The sample sizes were 20, 50, 200, and 400 responses. Second, for each sample, there were two penalty requirements. The sample was introduced with ±20% deviation from matching requirement, and then, after response proportions stabilized, the penalty zone was increased to ±40%. For example, consider the subject on the VI 30-s VI 30-s schedule when a 20-peck sample and a ±40% penalty zone were in effect. For a sample of 20, matching implies 10 responses to each key; 40% of 20 equals 8; thus right-response counts between 2 and 18 were ineligible for reinforcement. For the 2 birds at unequal left and right schedules, the penalty zone was calculated slightly differently when a ±40% requirement was in effect. We wanted to allow deviations to either side of matching. Thus, there was the added rule that right-response counts equal to the window (e.g., 20) and the window minus 1 (e.g., 19) were eligible for reinforcement (see Table 2). For example, with a 1:3 reinforcement ratio, a 20-peck win-

Table 2

Experiment 2: Order of conditions.

Subject	Condition	Window size	Deviation requirement (%)	Penalty zone
1	VI 30 s	20	20	$6 < r < 14$
	VI 30 s		40	$2 < r < 18$
		50	20	$15 < r < 35$
			40	$5 < r < 45$
		200	20	$60 < r < 140$
			40	$20 < r < 180$
		400	20	$120 < r < 280$
			40	$40 < r < 360$
2	VI 60 s	20	20	$11 < r < 19$
	VI 20 s		40	$7 < r < 19$
		50	20	$27 < r < 47$
			40	$17 < r < 49$
		200	20	$110 < r < 190$
			40	$70 < r < 199$
		400	20	$220 < r < 380$
			40	$140 < r < 399$
3	VI 150 s	20	20	$14 < r < 19$
	VI 16.6 s		40	$10 < r < 19$
		50	20	$35 < r < 49$
			40	$25 < r < 49$
		200	20	$140 < r < 199$
			40	$100 < r < 199$
		400	20	$280 < r < 399$
			40	$200 < r < 399$

dow, and a 40% penalty zone, right-response counts from 7 to 18 were ineligible for reinforcement. Third, for the purpose of data collection, choice proportions based on the moving average were calculated. For example, with each response the number of right responses in the last 20 was recorded. Fourth, and perhaps most important, during experimental sessions a correlated stimulus was added. The houselight color was linked to the distribution of choices. It was white when choice proportions exceeded the deviation from matching requirement, and it was blue when choice proportions failed to exceed the requirement. Thus, reinforcers were never delivered while the houselight was blue but could be delivered while the houselight was white. Each condition, meaning a sample size and deviation requirement, was kept in effect for 13 to 15 sessions.

RESULTS

Figures 5a, 5b, and 5c show the effect of the experimental contingencies on moment-to-moment choice proportions. On the *x* axis of each panel is the number of right responses out of the last *n* responses, where *n* is the window size. For example, the *x* axis goes from 0 to 50 for those sessions in which the reinforcement contingency was based on the just previous 50 responses. On the *y* axis is the frequency of each possible right-response count, as tallied after every response. For example, consider the expected right-response count frequencies for a session in which left and right responses occurred in strict alternation: left, right, left, right, and so on. The right-response count would be stuck at exactly one half the window size (e.g., 10 for a 20 window, or 200 for a 400 window) for the entire session, and as a result the histogram would have only one value: window size/2. In contrast, if responses did not occur in a strict alternating pattern, right-response counts would vary, and the corresponding histograms would reflect this. More generally, the shape of the histograms in Figure 5 reflects the pattern of switching between the two keys. The diagonal lines mark the choice proportions that were not eligible for reinforcement (penalty zone), and the single vertical line in the zone indicates the programmed reinforcement proportion. The data were taken from the last session that each condition was in effect. However, there was the capacity to record response counts for different window sizes simultaneously. (That is, one array was filled with right-response counts over the last 20 responses. Another array was filled with right-response counts over the last 50 responses, etc.) Consequently, for the baseline condition, different-sized histograms were from the same session, and differences reflect only the sampling process. The results from Pigeon 3 will be used to introduce some of the more important findings.

The average programmed interreinforcement intervals were 150 s and 16.6 s so that a 90% preference for the right key defined matching. When the window size was 50, the most frequent right-response count was the rightmost value on the *x* axis, 50. This means that Pigeon 3 often responded at least 50 times on the right key (VI 16.6 s) before switching to the left key (VI 150 s). However, when response counts were tallied in terms of samples of the just previous 400 responses (top right panel), the rightmost point on the *x* axis (i.e., right-response counts of 400) had a frequency of 0. This was because there were no instances in which there was a run of at least 400 responses in a row on the right side.

The second and third rows of each graph show the right-response count histograms when the penalty for matching was in effect. There were marked changes in both the shape and center of gravity of the distributions. The results for Pigeon 1 are representative. The programmed relative reinforcement frequency was 50% (VI 30 s VI 30 s), and in baseline sessions, response counts were usually within 20% of matching, even when calculated in terms of the 20-response window (top left panel of Figure 5a), implying that changeover rates were high. However, with the introduction of the 20% penalty, response proportions shifted from an even left-right distribution to a strong right-key preference. For example, most of the right-response counts fall outside the boundary of the 20% penalty zone. Similar results were obtained at the 40% penalty zone and at the larger averaging windows (e.g., 200 response averaging window). Thus, the overall right-response proportion shifted from about 50% to about 80% (whereas the right reinforcement

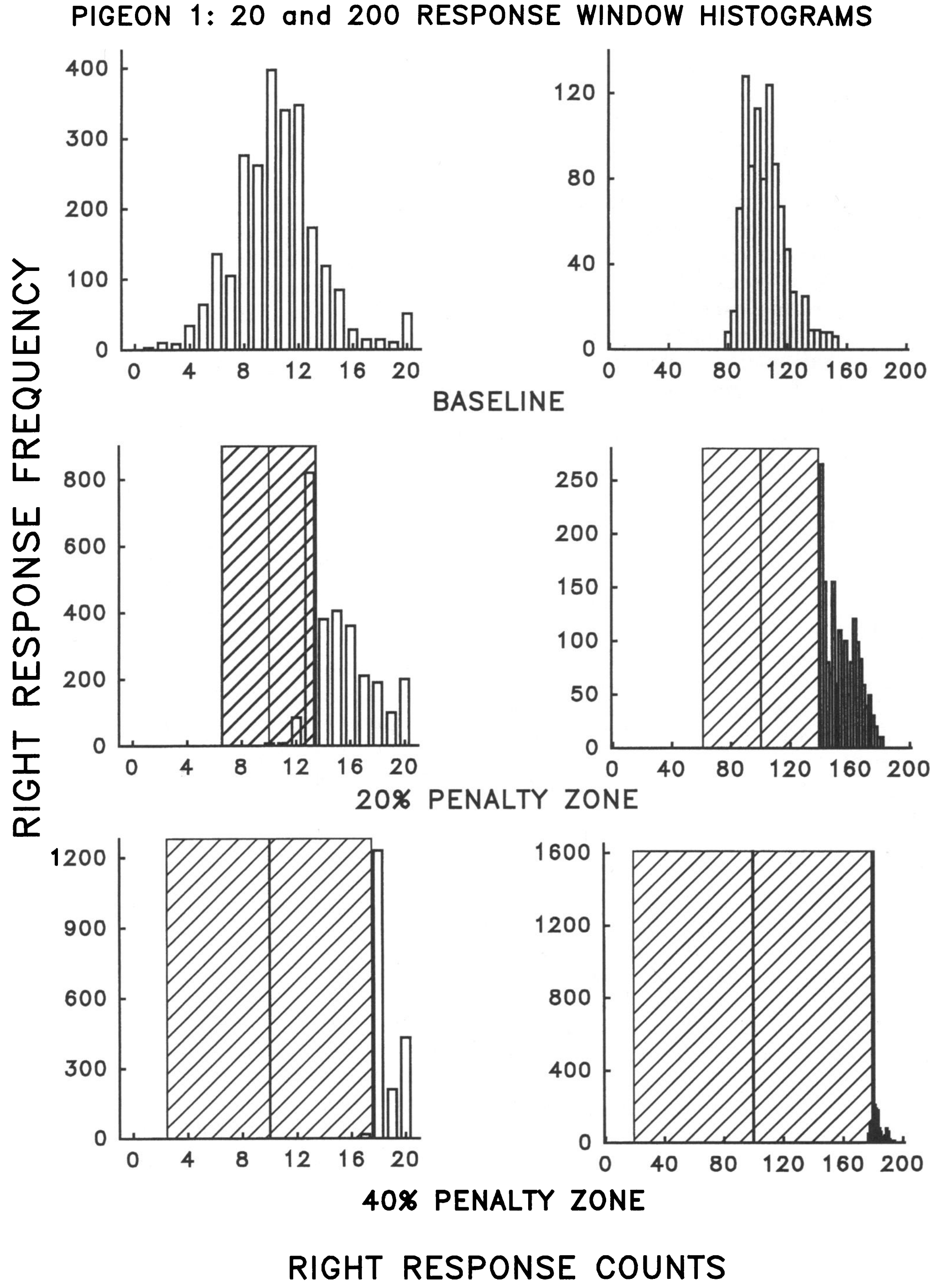

Fig. 5a. Right-response counts and the averaging window used to calculate deviations from matching for Pigeon 1. On the *x* axis is the number of right responses out of the last *n* responses, where *n* is the window size (20, 50, 200, or 400). On the *y* axis is the frequency of a particular right-response count. The top row shows baseline data (no deviation requirement); the second row shows effects of the 20% penalty zone; and the third row shows effects of the 40% penalty zone. The hatch lines circumscribe right-response counts that were not eligible for reinforcement. When response counts were within the hatch lines, the blue houselight was illuminated (see text). The data are from the last session of each condition.

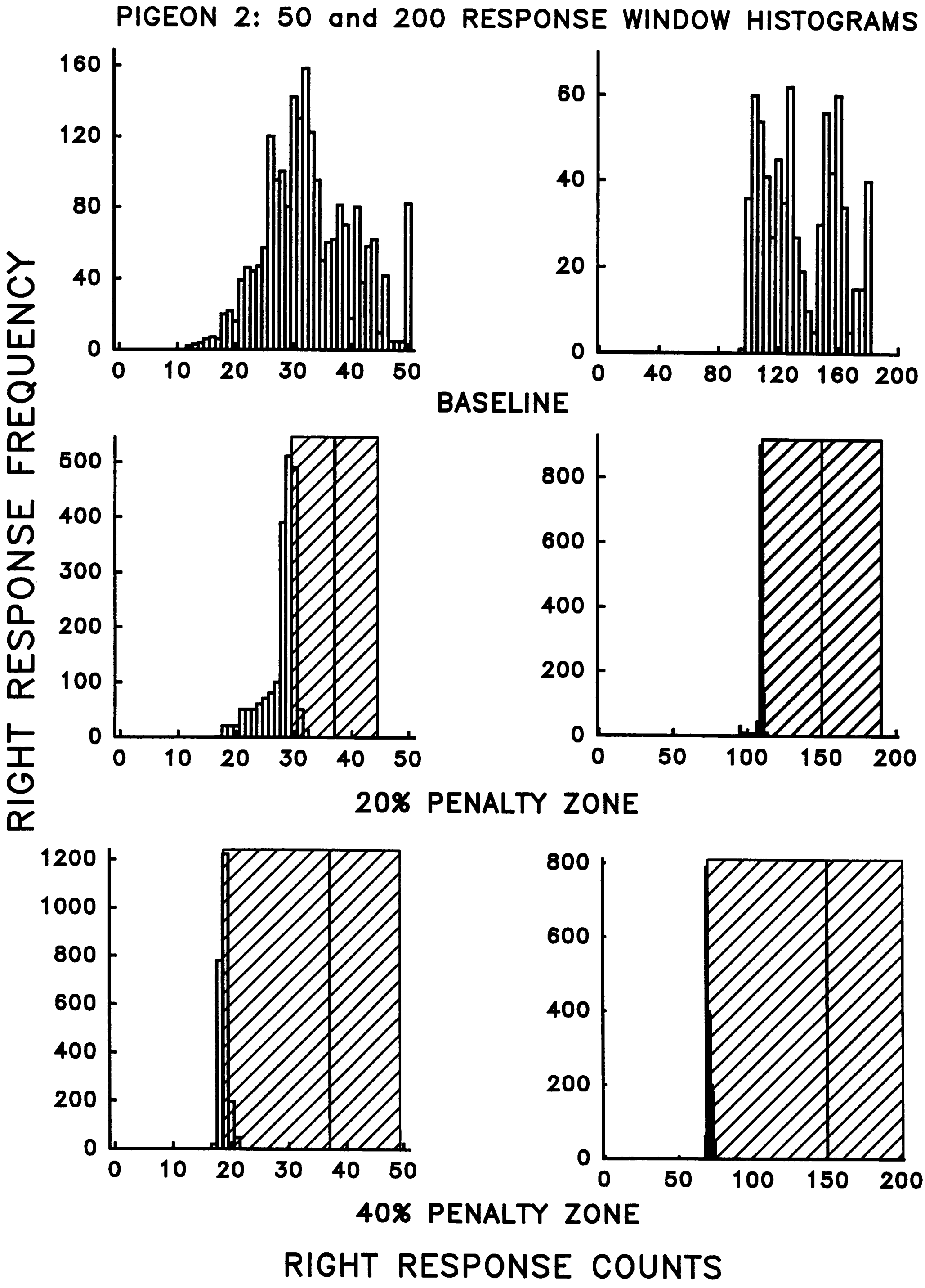

Fig. 5b. Right-response counts for Pigeon 2. The format is the same as in Figure 5a.

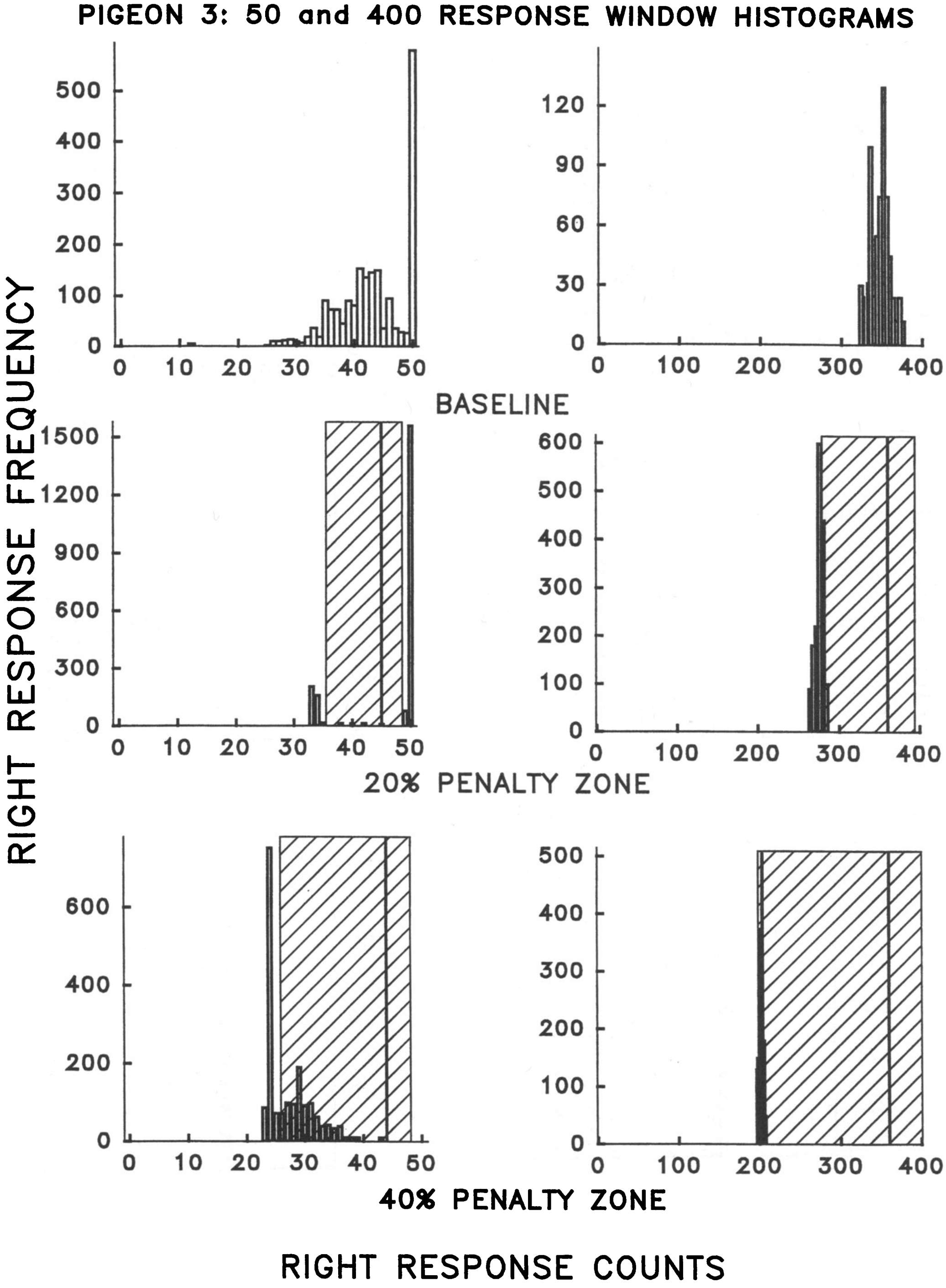

Fig. 5c. Right-response counts for Pigeon 3. The format is the same as in Figure 5a.

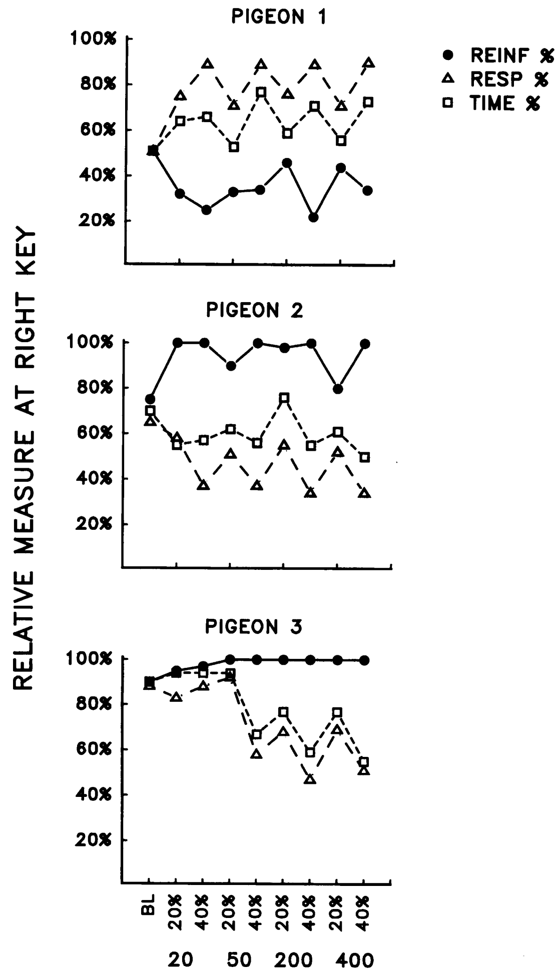

Fig. 6. Overall reinforcer, response, and time allocation on the right key in each condition. On the *x* axis is the experimental condition. The numbers 20 to 400 refer to the sample size used to calculate deviations from matching. The numbers 20% and 40% refer to the penalty zones for each sample size. The data are from the last session of each condition.

proportion shifted to about 30%; see Figure 6).

Figure 5a also shows that the penalty for matching also affected choice proportion variability. As the penalty zone increased, Pigeon 1's choice proportions became less variable. For example, the frequency of the modal right-response count markedly increased, and the width of the histograms greatly decreased. This pattern held for each subject and each averaging window.

The average programmed interreinforcement intervals for Pigeon 2 were VI 60 s VI 20 s. In baseline, right-response counts were centered at about 33 when the window size was 50 responses (implying an overall preference of about 66%). With the introduction of the experimental contingency, choice proportions shifted in the direction of the leaner schedule, to the very edge of the penalty zone. Consequently, when the penalty zone was large (40%), the shifts in response rates were large enough to result in overall preference for the leaner schedule (see Figure 6).

The effect of the contingency on Pigeon 3 is somewhat more complicated: Run lengths and response proportions changed. In baseline, overall choice proportions closely approximated the reinforcement proportions (perfect matching), and, unlike the other subjects, run lengths were quite long. For example, as discussed above, Pigeon 3 typically responded at least 50 times in a row on the right key (VI 16.6 s) before switching to the left key (VI 150 s). Consequently, when the 20-response and 50-response windows were in effect, choice proportions were usually 100% (20/20 and 50/50, respectively). This means that choice as measured in terms of the 20- or 50-peck window often exceeded the penalty zone, even though the overall (session-wide) choice proportion had not changed (and was within 2% of matching; see Figure 6). In support of these observations, in the first three experimental conditions, the frequency of long runs on the right key increased even further. For example, the number of right-response counts of 50 or more increased to over 1,500 when the averaging window was 50 responses and the penalty zone was ±20%. However, once the window size was large enough so that response counts were often within the penalty zone (40% penalty zone, 50-response window, Figure 5c), the distribution of choices shifted from the right key to the left key, as with Pigeon 2. For example, overall right choice proportions decreased to about 50% when the contingency was calculated in terms of a 400-response moving average. Thus, there were two ways to adapt to the contingency, and Pigeon 3 showed both as a function of the size of the penalty zone.

As measured by the shifts in response proportions, the experimental contingency was highly effective. In baseline, response proportions closely approximated reinforcement proportions; in experimental sessions, response proportions closely approximated the value predicted by the penalty zone. Moreover, it was possible to shape deviations from matching at every sample size. Thus, the houselight proved to be an effective discriminative stimulus.

The histograms (Figure 5) show the effects of the experimental contingency over small within-session samples of responding. In contrast, Figure 6 shows the effect of the experimental contingency at the session level. The data were collected from the last session of each of the nine conditions. (See Appendix B for time and response counts and changeovers.) In baseline, overall response and time proportions closely approximated reinforcement proportions (leftmost collection of points). However, under the experimental contingency, overall response proportions and reinforcement proportions diverged. For example, for the subject at equal left and right schedules (Pigeon 1), response proportions shifted to the right key, whereas reinforcement proportions shifted in favor of the left key. Similar results were obtained for Pigeons 2 and 3. In contrast to the results of Experiment 1, the time-allocation results followed the same pattern as did response allocation.

Figure 6 also shows that for Pigeons 2 and 3, reinforcement proportions were sometimes 100%, although response proportions had deviated toward 50%. This was a consequence of the dynamics of the contingency. For the 2 subjects on unequal left and right schedules, responding on the poorer side (the left) increased the likelihood that response counts would be eligible for reinforcement. However, it was more likely for a reinforcer to set up on the right side, and when this happened, the timer on the left side stopped. A switch to the right side would either earn or flush out the reinforcer and restart both timers. Eventually, a reinforcer would set up at the left side. However, if the subject had been responding on the right side sufficiently long, response counts would be within the penalty zone, so that when responding resumed on the left side, the primed reinforcer was flushed out rather than delivered. Thus, reinforcement proportions did not remain constant but became more extreme in favor of the better schedule, even though responding shifted in favor of the poorer schedule.

Figure 7 shows run length (average number of responses between changeovers) as a function of window size and penalty zone. The results for Pigeon 3 are most interesting. As shown in Figures 5 and 6, overall choice proportions remained near 90% (perfect matching) for the first three conditions of the experiment. Nevertheless, this subject lost fewer reinforcers than did Pigeons 1 and 2 (as shown in Figure 8). As described above, this occurred because of changes in run lengths. On both the left and right keys, the number of responses per visit to a side increased. Thus, when the window was relatively small, response counts usually exceeded the deviation requirement. However, when the penalty zone was increased to 40% in the 50-peck window, run lengths decreased and response proportions shifted away from matching.

Figure 8 shows the number of reinforcers obtained in the last session of each condition. There were two trends. There was a decrease in overall reinforcement rates in experimental sessions, and Pigeon 1 tended to earn more reinforcers when the penalty zone was 20% rather than 40%. However, in experimental sessions, overall reinforcement rate was about the same, regardless of window size. This was because response proportions typically exceeded the penalty zone.

DISCUSSION

In contrast to Experiment 1, there was no limit on deviations from matching. Moreover, in those cases in which relative reinforcement rate changed, relative response rate changed in the opposite direction. Also in contrast to Experiment 1, time allocation did not appear to be constrained by relative reinforcement rates. Given that the sample size was not a factor in Experiment 2, the major procedural difference between the two experiments was a discriminative stimulus for response proportions that were eligible for reinforcement rate.

In Experiment 2, reinforcement rate was maximized. The mediating processes may be molecular (e.g., the strengthening of sequences of left and right responses) or molar (e.g., the strengthening of a choice proportion). As described next, the data support a

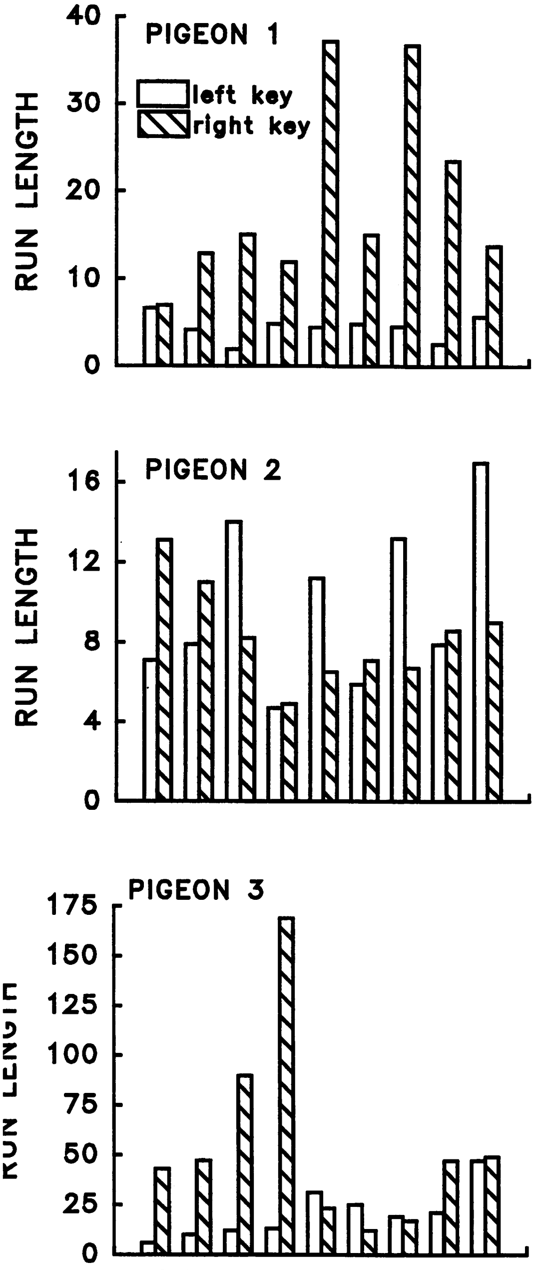

Fig. 7. Average number of responses between changeovers (run length). The numbers 20% and 40% refer to he penalty zones for each sample size. The data are from he last session of each condition.

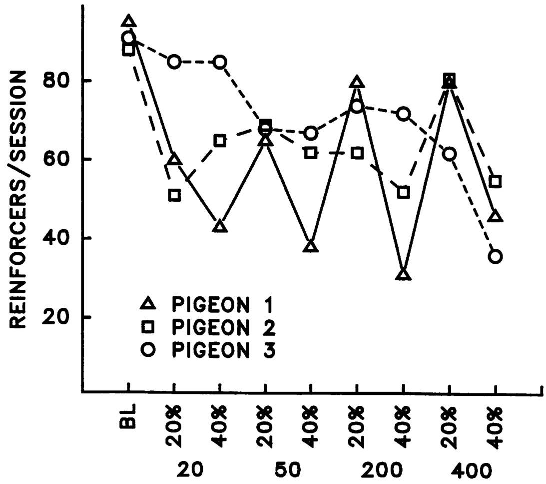

Fig. 8. Number of reinforcers per session. The numbers 20% and 40% refer to the penalty zones for each sample size. The data are from the last session of each condition.

molecular account. First, consider the idea that the deviation requirement strengthened specific response sequences or interreinforcement times (e.g., Shimp, 1975), but that in baseline, this sort of shaping process was less effective or nonexistent (e.g., Heyman, 1979; Real, 1983). The most immediate prediction is that the mean and variance of the response histograms should change. The mean should shift to near the edge or outside of the penalty zone, and the variance should decrease, especially in the large 200-response and 400-response windows. For example, if there was an invariant, repeating pattern of left and right responses (that was shorter than the averaging window), then the histogram would have but one entry, the number of right responses in the pattern. Figure 5 shows that the deviation contingency invariably reduced the variance in the histograms. Thus, there is evidence that maximizing was accompanied by an increase in sequential dependencies between responses.

An alternative hypothesis is that overall reinforcement rate more directly controlled choice. According to this view, choice proportions rather than specific response sequences were shaped. One implication of this theory is that responses should be more or less independent. If so then the expected standard deviation for the histograms would be the square root of *npq*, where *n* is the sample size and p and q are the proportions of right and left pecks (Heyman, 1979). For any given sample size (averaging window), this expression has a maximum at $p = q = .5$. However, Figures 5b and 5c show that just the opposite occurred. When preference moved towards 50%, the variance in the histograms decreased. Thus, it seems more reasonable to conclude that the processes that mediated maximizing in Experiment 2 operated at the level of response sequences or interchangeover times than at the level of response proportions.

Pigeons 2 and 3 lost relatively more reinforcers on the poorer side so that obtained reinforcement proportions were often 100% (at the richer side). Possibly this influenced the results. For example, did the pigeons maximize because in many sessions reinforcers were delivered only on one side? Pigeon 1's results put this question to rest. For Pigeon 1 the programmed relative reinforcement frequency was 50%, and in experimental conditions, relative reinforcement frequency never went to 100%. However, this subject adapted to the contingency as well as the 2 subjects on unequal left and right schedules did. Thus, maximizing reinforcement rate did not depend on the schedule values or the obtained reinforcement proportions.

Figure 6 shows overall choice and obtained reinforcement percentages. They differed markedly. However, the difference would be somewhat larger still if the comparison were restricted to just the period during which the white light was on. Recall that reinforcement could not occur while the current choice proportion, as defined by the averaging window, was within the penalty zone. That is, when the blue light was on, primed reinforcers were omitted rather than delivered. Figure 5 shows that this sometimes occurred. In Figure 6, response proportions are based on the session totals, including blue-light and white-light responding. If only white-light responding had been included, the difference between overall response and reinforcement proportions would have been slightly greater. (The reinforcement proportions are, of course, just for the white-light periods.)

GENERAL DISCUSSION

Experiment 2 differed from previous tests of economic explanations of matching in that discriminative stimuli identified changes in overall reinforcement rate. This simple procedural modification had a profound effect. In contrast to earlier and otherwise similar studies, relative reinforcement rate did not predict choice proportions. Rather, choice proportions were correlated with overall reinforcement rate. Put somewhat differently, rational choice theory predicted behavior when a discriminative stimulus signaled changes in overall reinforcement rate, whereas the matching law predicted behavior when location and key color signaled the component reinforcement rates (and there was no discriminative stimulus for overall rate).

The major theoretical challenge of this study is how to integrate the findings of Experiment 2 with previous studies in which relative rather than overall reinforcement rate predicted choice. One possible theory was outlined by Herrnstein (1982). He suggested that instances in which choice proportions were correlated with overall reinforcement rate would prove on examination not to violate matching. His point was that maximizing overall reinforcement entails the emergence of new response classes that, if properly measured, would obey the matching law. For example, as described above, the decrease in variability in the histograms implies an increase in response-by-response sequential dependencies. Possibly, these sequences of left and right responses were functional units, and the frequency of these new response units was proportional to their relative reinforcement rates. Rachlin, Green, and Tormey (1988) pointed out that similar arguments can be made to fit maximizing theory to apparently discrepant data. Unfortunately, the relations among maximizing, matching, and the dynamics of response-class formation have yet to receive much experimental attention.

However, whatever account of Experiment 2 eventually proves to be correct, it must have the flexibility to predict both matching and maximizing. For example, it seems reasonable to us to suppose the existence of an elementary reinforcement process that could yield matching or maximizing as a function of such factors as stimulus conditions. This perspective differs from the idea that individuals or species are either "maximizers" or "matchers." Instead, it suggests that reinforcement maximization and matching are outcomes (not processes) that depend on how the reinforcement contingencies are framed. Assuming this to be the case, the most immediate research questions are: What determines how contingencies are framed? And what is the underlying behavioral process that can lead to both matching and maximizing? The first question has practical as well as theoretical significance. For instance, the idea that contingencies can be framed so that preferences are either suboptimal or optimal has possible therapeutic applications (e.g., Herrnstein & Prelec, 1992; Heyman, in press). The second question is about scientific laws. Parsimony suggests that matching and maximizing are points along a continuum, generated by the interaction of a single underlying reinforcement process and situational factors such as stimulus conditions.

REFERENCES

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12,* 861–874.

Davison, M., & Alsop, B. (1991). Behavior-dependent reinforcer-rate changes in concurrent schedules: A further analysis. *Journal of the Experimental Analysis of Behavior, 56,* 1–19.

Davison, M., & Kerr, A. (1989). Sensitivity of time allocation to an overall reinforcer rate feedback function in concurrent interval schedules. *Journal of the Experimental Analysis of Behavior, 51,* 215–231.

De Carlo, L. T. (1985). Matching and maximizing with variable-time schedules. *Journal of the Experimental Analysis of Behavior, 43,* 75–81.

Findley, J. D. (1958). Preference and switching under concurrent scheduling. *Journal of the Experimental Analysis of Behavior, 1,* 123–144.

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., Rachlin, H., & Hanson, J. (1983). Matching and maximizing with concurrent ratio-interval schedules. *Journal of the Experimental Analysis of Behavior, 40,* 217–224.

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

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

Herrnstein, R. J. (1982). Melioration as behavioral dynamism. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 433–458). 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., Loewenstein, G. F., Prelec, D., & Vaughan, W. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making, 6,* 149–185.

Herrnstein, R. J., & Prelec, D. (1992). Melioration. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 235–263). New York: Sage.

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

Heyman, G. M. (in press). Resolving the contradictions of addiction. *Behavioral and Brain Sciences.*

Heyman, G. M., & Herrnstein, R. J. (1986). More on concurrent variable interval-ratio schedules: A replication and review. *Journal of the Experimental Analysis of Behavior, 46,* 331–351.

Hinson, J. M., & Staddon, J. E. R. (1983). Hill climbing by pigeons. *Journal of the Experimental Analysis of Behavior, 39,* 25–47.

Mazur, J. E. (1981). Optimization theory fails to predict performance of pigeons in a two-response situation. *Science, 214,* 823–825.

Rachlin, H., Green, L., & Tormey, B. (1988). Is there a decisive test between matching and maximizing? *Journal of the Experimental Analysis of Behavior, 50,* 113–123.

Real, P. G. (1983). A time-series analysis of change-over performance on concurrent variable-interval schedules. *Animal Learning & Behavior,* 255–265.

Savastano, H. I., & Fantino, E. (1994). Human choice in concurrent ratio-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 61,* 453–463.

Shimp, C. P. (1975). Perspectives on the behavioral unit: Choice behavior in animals. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 2, pp. 225–268). Hillsdale, NJ: Erlbaum.

Silberberg, A., & Ziriax, J. M. (1985). Molecular maximizing characterizes choice on Vaughan's (1981) procedure. *Journal of the Experimental Analysis of Behavior, 43,* 83–96.

Stubbs, D. A., & Pliskoff, S. S. (1969). Concurrent responding with fixed relative rate of reinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 887–895.

Vaughan, W., Jr. (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior, 36,* 141–149.

Williams, B. A. (1985). Choice behavior in a discrete trial concurrent VI-VR: A test of maximizing theories and matching. *Learning and Motivation, 16,* 423–443.

Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., Vol. 2, pp. 167–244). New York: Wiley.

Received January 23, 1995
Final acceptance July 9, 1995

APPENDIX A

Pigeon	Schedule	Condition (penalty zone)	Time (minutes)		Responses		Change-overs
			Left	Right	Left	Right	
45	VI 50 s	Baseline	20.5	15.8	1,960	2,033	228
	VI 50 s	10%	22.1	15.9	2,014	2,280	205
		15%	21.3	17.2	2,079	2,580	196
		20%	23.7	31.4	2,865	2,074	201
		25%	25.7	33.6	2,918	1,364	197
		30%	24.8	34.2	1,891	1,340	83
		35%	23.8	16.0	1,194	796	54
59	VI 33.3 s	Baseline	27.2	9.4	1,135	293	76
	VI 100 s	10%	28.4	9.5	948	341	105
		15%	26.7	10.6	560	320	101
		20%	25.6	12.1	867	603	126
		25%	28.4	11.1	632	474	77
		30%	28.6	11.3	359	239	37
62	VI 250 s	Baseline	3.3	32.7	360	2,232	50
	VI 27.8 s	10%	2.9	34.4	522	2,275	58
		15%	3.6	33.4	703	2,146	86
		20%	5.2	31.8	925	2,405	110
		25%	5.5	31.9	1,272	2,581	90
		30%	6.6	32.2	1,591	3,220	110
		35%	5.5	34.3	955	2,468	71
		40%	4.9	34.9	478	1,274	28
		45%	5.9	34.0	519	1,237	43
489	VI 33.3 s	Baseline	26.4	9.6	2,700	1,116	89
	VI 100 s	10%	27.5	9.7	2,492	1,153	137
		15%	27.3	10.0	2,798	1,790	141
		20%	27.0	11.3	2,604	1,685	167
		25%	26.3	12.5	2,328	1,785	162
		30%	26.3	13.2	2,670	2,054	157
		35%	27.6	12.1	2,223	1,548	116

APPENDIX B

			Deviation	Time (minutes)		Responses		Change-
Pigeon	Schedule	Window size	requirement	Left	Right	Left	Right	overs
1	VI 30 s	Baseline	—	12.4	12.9	1,237	1,339	376
	VI 30 s	20	20%	9.7	17.3	651	2,005	312
			40%	9.5	18.4	218	1,673	220
		50	20%	12.6	14.2	591	1,446	242
			40%	6.5	21.6	272	2,229	120
		200	20%	10.7	15.3	533	1,641	217
			40%	8.3	20.2	305	2,424	131
		400	20%	11.4	14.6	652	1,587	231
			40%	7.5	20.2	202	1,860	159
2	VI 60 s	Baseline	—	7.7	17.9	684	1,254	193
	VI 20 s	20	20%	12.4	15.1	889	1,238	226
			40%	11.5	15.3	1,162	671	165
		50	20%	10.1	16.5	878	922	375
			40%	11.8	15.1	1,432	844	357
		200	20%	9.4	17.5	847	1,014	286
			40%	12.3	15.1	1,455	736	220
		400	20%	10.1	15.9	913	999	232
			40%	13.6	13.6	1,363	709	159
3	VI 150 s	Baseline	—	2.6	22.9	229	1,718	80
	VI 16.6 s	20	20%	1.8	24.0	252	1,197	51
			40%	1.6	24.2	232	1,743	38
		50	20%	1.6	25.0	169	2,000	25
			40%	8.8	17.9	765	1,036	68
		200	20%	6.1	20.3	387	810	64
			40%	10.8	15.6	736	654	37
		400	20%	6.1	20.3	422	938	40
			40%	12.7	15.5	936	904	38