

## DECISION RULES AND SIGNAL DETECTABILITY IN A REINFORCEMENT-DENSITY DISCRIMINATION

MICHAEL L. COMMONS

COLUMBIA UNIVERSITY

Two probabilistic schedules of reinforcement, one richer in reinforcement, the other leaner, were overlapping stimuli to be discriminated in a choice situation. One of two schedules was in effect for 12 seconds. Then, during a 6-second choice period, the first left-key peck was reinforced if the richer schedule had been in effect, and the first right-key peck was reinforced if the leaner schedule had been in effect. The two schedule stimuli may be viewed as two binomial distributions of the number of reinforcement opportunities. Each schedule yielded different frequencies of 16 substimuli. Each substimulus had a particular type of outcome pattern for the 12 seconds during which a schedule was in effect, and consisted of four consecutive light-cued 3-second T-cycles, each having 0 or 1 reinforced center-key pecks. Substimuli therefore contained 0 to 4 reinforcers. On any 3-second cycle, the first center-key peck darkened that key and was reinforced with probability .75 or .25 in the richer or leaner schedules, respectively. In terms of the theory of signal detection, detectability neared the maximum possible  $d'$  for all four pigeons. Left-key peck probability increased when number of reinforcers in a substimulus increased, when these occurred closer to choice, or when pellets were larger for correct left-key pecks than for correct right-key pecks. Averaged over different temporal patterns of reinforcement in a substimulus, substimuli with the same number of reinforcers produced choice probabilities that matched relative expected payoff rather than maximized one alternative.

*Key words:* discrimination, decision rules, reinforcement, memory, matching, maximizing, signal detection, T-schedule, key peck, pigeon

Most studies of reinforcement schedules have been concerned with the effect of reinforcement schedules on responding maintained by these schedules. The present approach is to examine discriminative properties of probabilistic reinforcement schedules in a choice procedure (Estes, Burke, Atkinson, & Frankmann, 1957; Shimp, 1973). Psychophysical techniques then permit an analysis of stimulus control by reinforcement schedules. Rilling and MacDiarmid (1965), Pliskoff and Goldiamond (1966), Hobson (1975), and Lattal (1975) used schedules of reinforcement

where the number of responses per reinforcer were the discriminative stimuli. In the present study, reinforcement density provided the basis for discrimination.

Using a signal detection analysis, Rilling and MacDiarmid (1965) studied how pigeons discriminated fixed ratio (FR) reinforcement schedules. When a center key came on, one of two FR schedules was arranged for responding on it. Meeting the center-key requirement darkened that key and turned on two side keys. After the lower valued FR, which may be considered the signal plus noise ( $S + N$ ), a left-key peck was reinforced and a right-key peck darkened the chamber. After higher valued FR, which may be considered noise ( $N$ ), a right-key peck was reinforced and a left-key peck darkened the chamber. The discriminability of the signal plus noise ( $S + N$ ) schedules from the noise ( $N$ ) schedules was measured from isosensitivity curves where the coordinates were the conditional probabilities of left-key pecks given signal plus noise,  $P(\text{Hit}) = P(L|S + N)$ , and of left-key pecks given noise alone,  $P(\text{False Alarm}) = p(L|N)$ . These choice

---

This research was supported in part by Dare Association Grant #103. Portions of the data were reported at the 1973 Eastern Psychological Association convention, and served to partially fulfill the requirements for a PhD degree. The author wishes to thank John Anthony Nevin and Bruce Schneider for their helpful supervision; Patrice M. Miller, Thomas R. Farrell, and Peter Pastor for assistance in conducting the experiment; and my numerous colleagues for editorial suggestions. Reprints may be obtained from Michael L. Commons, Laboratory of Human Development, Harvard Graduate School of Education, Harvard University, Larsen Hall, Appian Way, Cambridge, Massachusetts 02138.

probabilities provided a sensitive index of the pigeon's discrimination of the schedules.

Since a major difference between reinforcement schedules is their reinforcement density, and since that difference has been shown to control rates and probabilities of responding (Catania and Reynolds, 1968; Herrnstein, 1970), the present study examined the discriminative control exercised by different reinforcement densities.

The discriminability of reinforcement density was examined in two ways. First, isosensitivity (ROC) curves were determined to see how well pigeons discriminated schedules differing in their reinforcement probability,  $p$ , and corresponding reinforcement density,  $N_d$ . A discriminability index,  $d'$ , was found for the actual birds by comparing the obtained data to points and curves derived from assuming various models of optimizing ideal observer performance (an observer that maximizes obtained reinforcement).

Second, decision rules were estimated from the data and compared to an ideal observer's decision rules. A decision rule is defined as the conditional probability of a particular choice given that certain situations or parameters of situations have occurred. It describes the control of choice by various parameters of a stimulus, such as reinforcement amount and proximity of reinforcement opportunity to choice. It also provides a framework in terms of which various versions of maximization laws and matching laws may be examined.

## METHOD

### Subjects

Four male White Carneaux pigeons with extensive multiple-schedule histories served. The birds were maintained at approximately 80% of their free-feeding weights throughout the experiment.

### Apparatus

Two standard Lehigh Valley pigeon test chambers with special 1519B Pigeon Intelligence Panels were used. Three translucent Lehigh Valley pecking keys, 2.54 cm in diameter, were spaced 8.25 cm apart center to center, with the center of the outer keys 9.21 cm from the chamber sides. All keys were 25.40 cm from the floor. A model 111-05 stimulus light transilluminated each key. A peck of at least 9.0 g

(.9 N) operated a key's microswitch. The key colors were red, green, and yellowish white for the left, right, and center, respectively.

The bottom of a feeder aperture (an opening 5.72 cm wide and 5.08 cm deep) was centered 10.0 cm above the floor. Two pellet feeders in each box dispensed pellets into the hopper. One feeder delivered 20-mg pellets after designated center-key pecks, and after correct side-key pecks under some conditions. The second feeder delivered either 45-mg pellets or 97-mg pellets, depending on the condition. The sounds of the different pellets as they fell into the hopper were discriminably different to a human observer. The hopper light was illuminated for 2.7 sec upon feeder operation.

A houselight located above the center key was illuminated during most of the procedure except after a peck at an incorrect key during pretraining.

### Stimuli

The two stimuli to be discriminated in the present situation may be viewed from two different perspectives. One emphasizes that each stimulus is a schedule of reinforcement; the other looks at each stimulus as a distribution of the number of reinforcement opportunities programmed on a center key.

From the first view, two T-schedules of reinforcement, one rich in reinforcement,  $S_{rich}$ , the other lean,  $S_{lean}$  were overlapping stimuli to be discriminated in a choice situation. A T-schedule (Schoenfeld, Cumming, & Hearst, 1956) consists of a T-sec cycle which contains a subcycle,  $t^D$ , the time period when behavior may be reinforced with some probability,  $p = p(S^R + |R)$ . The T-schedule modifications used here were similar to Weissman's (1961) changes, particularly his procedure of cueing the beginning of the  $t^D$  portion of a 90-sec T-cycle by illuminating the key. With this procedure, only the first response in  $t^D$  produces a reinforcer, with probability  $p(S^R + |R) = 1$ . In the present experiment, each trial's 12-sec stimulus period contained four 3-sec cycles during which a substimulus sampled from one of the two  $p$ -valued T-schedule stimuli was presented.

From the other view, these two schedule stimuli can be considered as consisting of binomial distributions of the number of reinforcement opportunities in the sampled sub-

stimulus, with each of the two schedule stimuli yielding different frequencies of the same 16 possible substimuli, as shown in both Figure 1 and Table 1. As noted above, each substimulus sample consisted of four consecutive light-cued 3-sec cycles ( $t^D = T = 3$ ). On each 3-sec cycle, the first center-key peck darkened that key and could be followed by reinforcement. The probability of reinforcement for a center-key peck was  $p(S^{R+}|R_C) = .75$  when the rich schedule was in effect and .25 when the lean schedule was in effect. If a bird failed to peck the key in 3 sec, food was not delivered and the next 3-sec cycle was immediately initiated. Each set of four cycles will be referred to as a substimulus sample,  $S_n$ . The pigeon's task was to discriminate whether the rich or the lean schedule was in effect during this sample of four cycles.

There are four levels of description of the stimuli.

1. On a "gross molar" or gross level, each probabilistic reinforcement schedule, either the

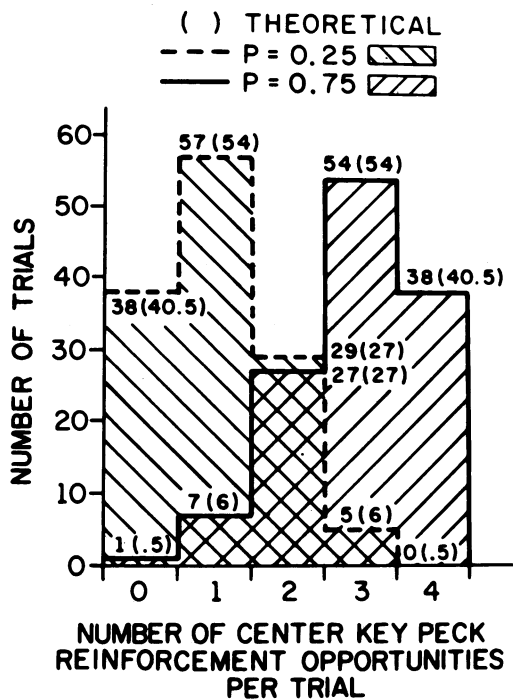


Fig. 1. Two sets of binomial frequency distributions show the actual rich and lean stimulus reinforcement distributions for center-key-pecks. The substimulus densities, the number of possible reinforcement opportunities in four cycles, ranged from 0 to 4. On the rich schedule the probability of reinforcement for the first center peck in a cycle,  $p$ , was .75, and on the lean  $p$  was .25.

rich ( $p = .75$ ) or the lean ( $p = .25$ ) is viewed as a single "stimulus", here called  $S_{rich}$  and  $S_{lean}$  respectively: sampling considerations are of no concern.

2. At the "molar" level, all substimulus samples with the same number of reinforcements are viewed as the same "stimulus". Each substimulus,  $S_n$ , has a reinforcement density,  $D(S_n)$ , equal to  $N_d$ , the number of center-key peck reinforcement opportunities over the four cycles in the substimulus. Density, the defining characteristic of a "stimulus" at this level of analysis, ranges from zero to four reinforcements per substimulus, giving rise to five such "stimuli". The same-density substimulus,  $S_{N_d}$ , is the general member of a given "stimulus" set, with density  $N_d$ . This level of analysis distinguishes between number of reinforcements within a sample stimulus but not between the particular 3-sec cycles within that 12-sec stimulus period on which those reinforcements are programmed.

3. At the "molecular" level, the definition of a "stimulus" involves the pattern of reinforcement in a substimulus. Each substimulus is represented as a four-digit binary number. A number such as 0001 indicates a sample with three cycles without reinforcement opportunities followed by one cycle with an opportunity. The substimuli,  $S_n$ , are numbered from 0000 to 1111. The left-most digit represents the cycle furthest in time from choice, and the right-most represents the cycle immediately before choice.

4. At the "micro" level, the definition of a "stimulus" depends on whether or not there is a reinforcement opportunity on a particular cycle,  $t$ -seconds preceding choice, irrespective of what is programmed for its neighbors.

There are 16 possible combinations in a 4-cycle substimulus. These are shown in Table 1 along with the probability that the particular substimulus will occur given that the rich (.75) or lean (.25) schedule was presented. The probability that a particular set of four events will occur was obtained by expanding the binomial  $(p + q)^4$  with  $p = .75$  or  $.25$  and  $q = 1 - p$ . Actual presentation frequencies deviated somewhat from the expected frequencies because of sampling. Therefore, both the expected and actual frequencies of each event across the 256 trials in a session are given in the table.

Table 1

Description of stimuli. For each of the five possible molar classifications of substimuli with number of reinforced cycles,  $N_s$ , going from 0 to 4, and each of the 16 possible molecular substimuli,  $S_n$ , going from 0000 to 1111, the theoretical and sampled probabilities of reinforcement are given for the rich (LC),  $p = .75$ , and the lean (RC),  $p = .25$ , schedules. The sum of the molecular frequencies equals the molar frequencies. On a molecular level 0 indicates a nonreinforced cycle and 1 a reinforced cycle, the digit on the right being closest to choice. The theoretical and sampled numbers of occurrences of molar classified substimuli and molecular substimuli in the 256 trials from a session are also given.

Molar Level: Number of reinforcers in a substimulus $N_s$	Molecular Level: Substimulus number in binary notation	Gross Molar					
		Rich frequency (LC)		Lean frequency (RC)		Frequency of Corrects (LC + RC)	
		Theo- retical	Sam- pled	Theo- retical	Sam- pled	Theo- retical	Sam- pled
0	0000	0.5	1	40.5	38	41	39
		0.5	1	40.5	38	41	39
	1000	1.5	1	13.5	17	15	18
	0100	1.5	1	13.5	16	15	17
	0010	1.5	3	13.5	14	15	17
1	0001	1.5	2	13.5	10	15	12
		6.0	7	54.0	57	60	64
	1100	4.5	3	4.5	4	9	7
	1010	4.5	5	4.5	6	9	11
	1001	4.5	6	4.5	4	9	10
	0110	4.5	3	4.5	5	9	8
	0101	4.5	6	4.5	6	9	12
	0011	4.5	4	4.5	4	9	8
2		27.0	27	27.0	29	54	56
	1110	13.5	17	1.5	2	15	19
	1101	13.5	14	1.5	1	15	15
	1011	13.5	11	1.5	0	15	11
	0111	13.5	12	1.5	2	15	14
3		54.0	54	6.0	5	60	59
	1111	40.5	38	0.5	0	41	38
4		40.5	38	0.5	0	41	38
		128	127	128	129	256	256

A molecular-level stimulus, even 0000 or 1111, could occur with either schedule stimulus. However, the greater the number of cycles having a reinforcement opportunity, the greater the likelihood that the rich schedule was in effect. Therefore, the lean (or rich) schedule was more likely to be in effect when 0 or 1 (or 3 or 4) reinforcement opportunities were presented on a trial (respectively). Substimuli with 2 reinforcement opportunities occurred about equally often given either schedule.

The procedure used here may be interpreted not only in terms of four cycles of a cued T-schedule but also in terms of a cued random-interval (RI) or variable-interval (VI) schedule

with an average interval length equal to 3 sec times  $p(S^{R+}|R_C)$ , which is  $\frac{3}{4}$  sec and  $\frac{3}{4}$  sec for the lean and rich stimuli, respectively. The T-schedule in some ways resembles a VI schedule, and therefore the results from the present experiment may help to interpret multiple and concurrent VI-VI schedule performance. For the organism, a cued T-schedule differs from clocked VI (in which the beginning of an interval is timed from the end of the first interval, irrespective of the organism's behavior) in two ways: (a) the beginning of every cycle (interval) was cued and (b) the first response in the first cycle (interval) could be reinforced. The cue light that began the clocked cycle and the reinforcement schedule brought

the birds near the hopper and insured a short latency center-key peck at the beginning of each cycle. This presumably insured the observation of reinforcement delivery. Clock-determined cycle length together with constant reinforcement probability within each cycle provided a precise a priori description of a trial's four-cycle reinforcement pattern and a uniquely defined substimulus.

A number of possible distributions could have been used to program a substimulus' reinforcement opportunities during the trial's stimulus period. Rilling and MacDiarmid (1965), Pliskoff and Goldiamond (1966), and Hobson (1975), used nonoverlapping distributions of response events. Because decision rules that reflect how various aspects of the gross schedule stimuli control choice were of interest here, it was important to have the discrimination difficult enough to guarantee that the birds would make a number of errors. To insure this, there had to be trials where the particular temporal sequence of reinforcement and nonreinforcement opportunities could have come from either schedule stimulus distribution. This overlap of the two well-defined stimulus distributions could allow some of the notions from signal detection theory to be applied and examined. On a trial, for each cycle within a substimulus sample from a stimulus distribution, the first key peck had to have the same reinforcement probability. Hence, binomial distributions,  $\beta(N,p)$  and  $\beta(N,1-p)$ , with  $N = 4$  (cycles), lean schedule  $p$  value of .25 and rich schedule  $p$  value of .75 were dictated.

### Procedure

The experiment was divided into two stages, pretraining and training. Sessions were 76.8 min and 256 trials long. During training, the stimuli to be discriminated were the lean and rich reinforcement schedules described above. During pretraining, however, the probability of reinforcement for a center-key peck for the rich schedule was 1.0 and for the lean schedule was 0. During the first step of pretraining, the substimulus consisted of two 3-sec cycles. Within a trial, each cycle's first key peck was reinforced with the same probability as all other cycles in that trial and always darkened the center key and, at first, also the house-light, both of which were reilluminated at the start of the next cycle. If no pecks occurred the

light(s) remained on for 6 sec and no reinforcement was delivered.

On completion of the 12-sec stimulus presentation, a 6-sec choice period began with the illumination of the two side keys, while the center key either stayed dark or was darkened. For a rich schedule substimulus, the first left-key peck was reinforced ("Left Correct," or LC) and a right-key peck was unreinforced ("Right Error," or RE). For a lean schedule substimulus, the first right-key peck was reinforced ("Right Correct," or RC) and a left-key peck was unreinforced ("Left Error," or LE). The first side-key peck darkened and deactivated both keys until the choice period ended.

The second step of pretraining began when at least 95% of the choices were correct for two successive sessions. Each substimulus now consisted of four cycles. The probability that a center-key peck would be followed by reinforcement remained at 1.0 when the rich schedule was in effect and at 0 when the lean schedule was in effect. Corrective training (given as needed) for position preferences was to present the rich schedule continually for excessive right-key pecking and the lean schedule for excessive left-key pecking. Together, the two pretraining steps took up to 30 sessions.

The four birds were then trained to discriminate between the probabilistic rich ( $p = .75$ ) and lean ( $p = .25$ ) schedules. Each bird was trained under different biasing conditions until discrimination performance was stable. These conditions differed in terms of size of the pellet obtained for correct choices of the left and right keys: For a correct left-key peck, LC, pellet sizes were 97 mg, 45 mg, 20 mg, 20 mg, 20 mg, and for a correct right-key peck, RC, they were 20 mg, 20 mg, 20 mg, 45 mg, and 97 mg, for Conditions 1, 2, 3, 4, and 5, respectively. Thus the ratios of pellet sizes for correct left-key to correct right-key pecks under the five conditions were, in order, 5:1, 2.25:1, 1:1, 1:2.25, 1:5. The corresponding relative amounts of reinforcement for correct left-key pecks were .833, .692, .500, .308, and .167.

Data were obtained at each biasing condition after choice behavior stabilized. Stabilization was said to have occurred when the direction of change in left-key peck probability,  $p(L)$ , had reversed itself at least twice after a session, and the average of the next two sessions' change was less than  $\pm .05$ . Data were

first collected following training with pellets of equal size (20 mg) for both correct left-key and correct right-key pecks (Biasing Condition 3). Then half the birds were given a 97-mg pellet for correct left-key pecks and a 20-mg pellet for correct right-key pecks (Biasing Condition 1), and the other half were given a 45-mg pellet for correct left-key pecks and a 20-mg pellet for correct right-key pecks (Biasing Condition 2). Then the birds in Biasing Condition 1 were switched to Biasing Condition 2, and the birds in Biasing Condition 2 were switched to Biasing Condition 1. Next, the conditions were reversed with the choice now biased to the right; half the birds were given a 20-mg pellet for correct left-key pecks and a 97-mg pellet for correct right-key pecks (Biasing Condition 5), and half were given a 20-mg pellet for correct left-key pecks and a 45-mg pellet for correct right-key pecks (Biasing Condition 4). Then Biasing Conditions 4 and 5 were switched for each group of birds.

## RESULTS

Figure 2 shows the sensitivity data of Birds 27, 29, 31, and 85. The probability of a left-key peck given that a rich schedule sample was presented,  $P(\text{Hit})$ , is shown as a function of the probability of a left-key peck given that the lean schedule sample was presented,  $P(\text{False Alarm})$ . The x's on the solid isosensitivity curves, also called receiver operating characteristic (ROC) curves, are plots of theoretical probabilities of hits versus false alarms expected for the distributions shown in Figure 1, given (in a continuous model described below) three different levels of sensitivity, or  $d'$ . These x's would be obtained if a pigeon followed a decision rule described by statistical decision theory. A particular integral value of substimulus reinforcement density serves as a criterion. For substimuli perceived to have a reinforcement density above this criterion only left-key pecks are made; for substimuli below, only right-key pecks are made.

These three isosensitivity functions are found by determining  $P(\text{Hit})$  and  $P(\text{False Alarm})$  coordinates from an appropriate Lebesgue integral of the binomial distributions in Figure 1, on the assumption that the means of the rich and lean stimuli were separated by 2.0, 1.5, and 1.0 reinforcement opportunities (maximal to moderate separation). For the bi-

nomial distributions in Figure 1, criteria were set at each boundary where the number of reinforcement opportunities incremented. These criteria were used to establish the points indicated by x's shown in Figure 2.

To test whether or not birds in some sense "counted" the number of reinforcements in a substimulus, performance was compared to predictions obtained from an ideal observer who counted and whose performance depended on an assumed discrete distribution. Performance was compared also to predictions obtained from an ideal observer who did not count whose performance depended on an assumed pointwise discontinuous distribution. The details of the models are presented in the Appendix. Two versions of each model are considered. One assumes that only the difference between the mean number of reinforcements for the rich and lean schedule decreases (while the variance remains constant), and the other assumes that the differences between probability of reinforcement for the rich and lean schedules decrease, producing a decrease in separation between means. These two forms of the two models are presented so that the  $d'$  values corresponding to the theoretical points or curves may be more clearly interpreted and more nearly comparable to results from other experiments. The actual data points did not cluster around the points predicted by either form of the discrete counting model, suggesting that the birds did not, in the sense defined by these models, "count" reinforcements.

At the gross molar level, if the birds followed the statistical decision rule described earlier and correctly identified each substimulus density, the probability of hits and false alarms obtained would fall on the top isosensitivity curve. Points below this line indicate that performance is worse than that of the second ideal observer ( $d' = 2.19$ ). Biases introduced by changing relative amount of reinforcement for correct left-key and correct right-key pecks should appear as different points on the same isosensitivity curve.

Greater sensitivity is displayed in Figure 2 by empirical points lying closer to the top one of the curves for ideal observers. Point clustering indicates a lack of day-to-day variation. The degree of the sensitivity for the four birds, in order from least to most sensitive bird, is 85, 27, 29, and 31, with the sensitivities of

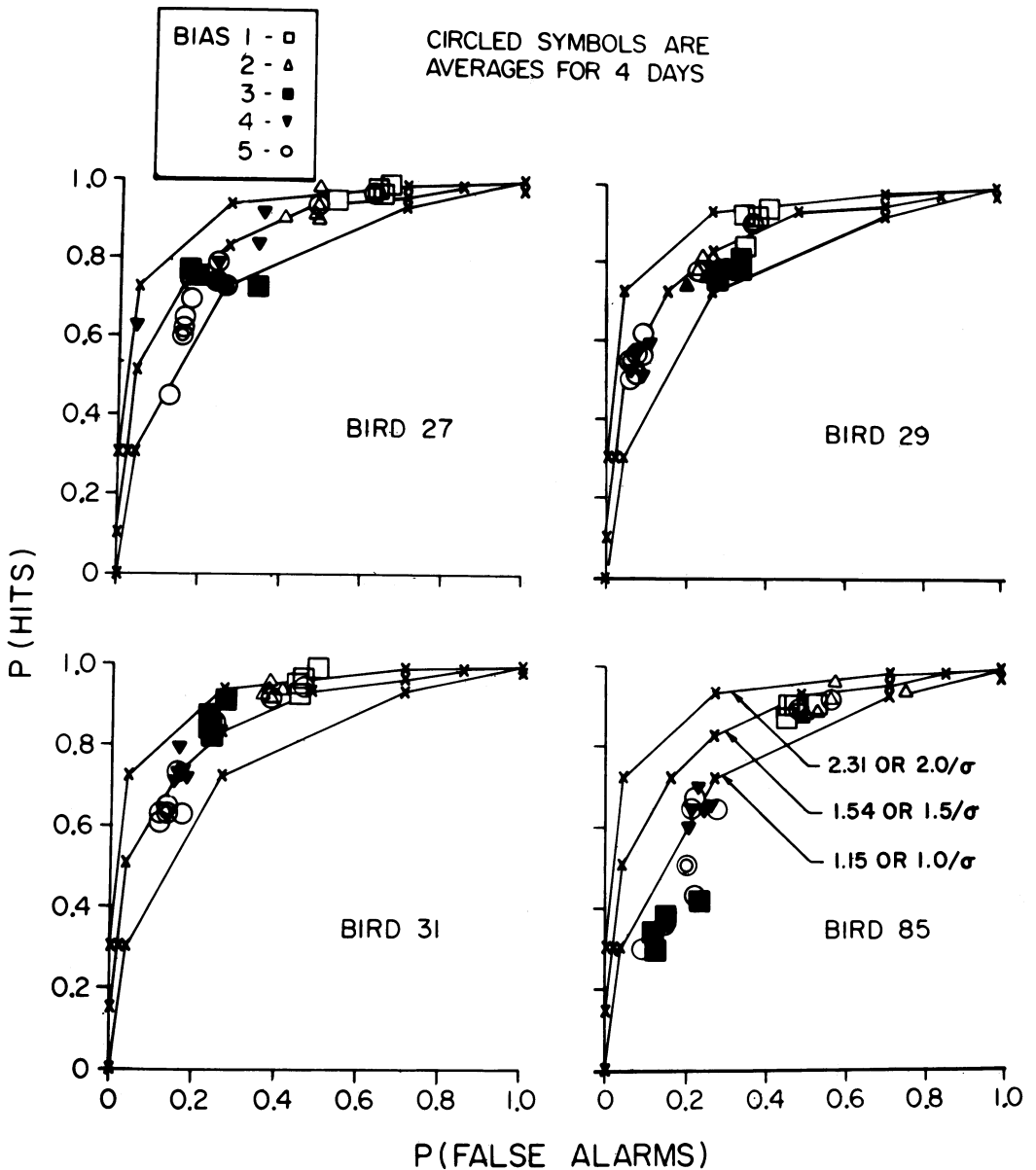


Fig. 2. Isosensitivity curves. The probability of a Hit,  $P(\text{Hit})$ , is equal to the number of correct left-key pecks divided by the number of possible correct left-key pecks. The probability of a False Alarm,  $P(\text{False Alarm})$ , is equal to the number of incorrect left-key pecks divided by the number of possible correct right-key pecks. The solid lines show isosensitivity curves for three values of  $d'$ . The top curve would be obtained if the subject followed a second ideal decision rule (see text).

Birds 27 and 29 being similar. The values of  $d'$  for many of the conditions were usually between the top and bottom isosensitivity curves, except for Bird 85 at Biasing Conditions 3 and 5. The biasing effect was least for Bird 31.

The difference between maximum attainable performance and the obtained perfor-

mance might be attributable to normally distributed random errors. Later information will show that this assumption is not appropriate here, insofar as it is possible to obtain representations of the decision rules and to find nonnormally distributed error sources, such as the effect of the decrease in stimulus control as a function of increasing time and

therefore decreasing memory for important events.

The effects of three experimental variables that may control choice will be analyzed: (a) the relative reinforcement amount for a correct left-key peck, (b) the number of reinforcement opportunities in a substimulus, and (c) reinforcement opportunity proximity in time to the choice point.

Bias, which is the probability of a left-key peck,  $p(L) = (LC + LE)/(LC + LE + RE + RC)$ , was examined as a function of biasing condition in Figure 3. Relative Reinforcement or biasing,  $A_{\beta L}$ , is the ratio of pellet size delivered for a correct left-key peck to the sum of left-correct pellet size and right-correct pellet size. Bias, or  $p(L)$ , is an increasing function of increasing  $A_{\beta L}$ : as shown in Figure 3, the left-key peck probability ranged from a  $p(L)$  of about .31 with a relative reinforcement amount of .17 to a  $p(L)$  of about .80 with a relative reinforcement amount of .83.

Bias did not differ substantially from bird to bird. Bird 85's one anomalous point might be attributable to its data being from the first condition conducted (Biasing Condition 3). This one point corresponds to a loss of substantial reinforcement as seen by the low  $d'$  in Figure 2. After exposure to other biasing con-

ditions, the behavior of this bird was similar to that of the others.

Recall from above that substimuli may be described on three other levels. On the molar level, only reinforcement density is considered and substimuli take on reinforcement density values from 0 to 4. On a molecular level, there are 16 substimuli, each indicating a particular reinforcement pattern that may occur over the 4 cycles. On the molecular level, substimuli are described using binary notation. For example, substimuli with density value of 1 at the molar level can have molecular values of 0001, 0010, 0100, and 1000.

Figure 4 shows both molecular and molar level data. At the molecular level, each row shows one bird's  $p(L)$  for each substimulus density, during the four final sessions of training on that condition. Each point indicates the probability of a left-key peck to a given substimulus,  $p(L|S_n)$ , under the given condition. At the molar level, for each biasing value, the decision rules are indicated by the average left-peck probability for substimuli at a given density,  $p(L|S_{N_d})$ . This weighted average was found by summing  $p(L|S_n)$  for each same-density substimulus, weighted by the frequency of the substimulus,  $p(L|S_{N_d}) = \sum_{D(S_n) = N_d} p(L|S_n)p(S_n)$ .

The empirical fact that the molecular points do not fall on the line through the weighted average  $p(L|S_{N_d})$  suggests that, although the substimuli are equivalent on a molar level by definition, they are not equivalent on a molecular level. That is, at the molecular level, substimuli act differently in determining the left-key peck probability. The proximity of reinforcement to choice appears to have a substantial impact. This will be considered next.

The effect of the proximity to choice of a reinforcement opportunity will be shown first, followed by the effect of the proximity to choice of the omission of a reinforcement opportunity. At the molar level, density 1 substimuli,  $D(S_n) = N_d = 1$ , provide exactly one reinforcement opportunity, and density 3 substimuli provide exactly 3 reinforcement opportunities. For the density 3 substimuli, one of the four cycles has a reinforcement opportunity omitted, making density 1 and 3 substimuli comparable opposites.

Figure 5 shows the effects of the distance between a choice and a single reinforcement

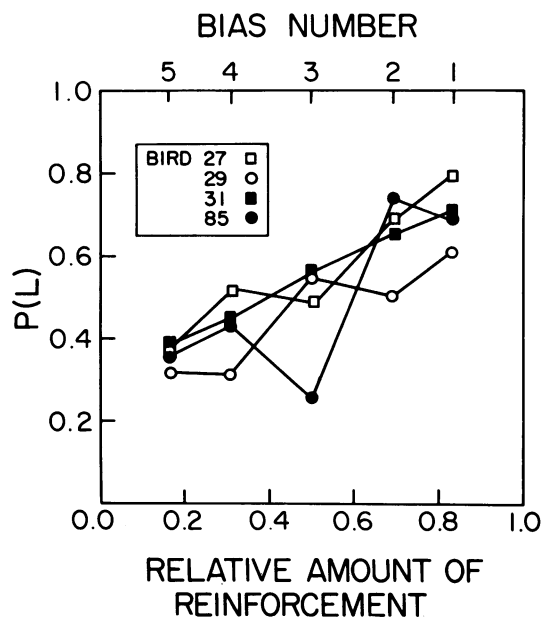


Fig. 3. The probability of a left-key peck,  $p(L)$ , plotted against the relative amount of reinforcement,  $A_{\beta L}$  (biasing), for correct left-key pecks.



Table 2

Expected molar-level decisions to peck left to a substimulus with a reinforcement density of  $N_d$ . The predictions of  $p(L|S_{N_d})$ , based upon matching the probability of a left-key peck for a given density substimulus, to the relative expected payoff for that choice, are shown as a function of the number of cycles in a substimulus containing a reinforcement opportunity ( $N_d$ ) and the relative reinforcement amount for correct left-key pecks (Biasing Condition). Left correct probability to a given density substimulus,  $p(L|S_{N_d})$  times the relative reinforcement amount,  $A_{\beta L}$ , equals amount of reinforcement for correct left peck relative to amount for left correct pecks and right correct pecks, yielding expected payoff.

Number of reinforcers in a substimulus $N_d$	Biasing	Expected $p(L)$ * at the Molar Level					Relative expected payoff for a left peck Rel EP(L  $S_{N_d}$ )
		Probability of a L or R being correct $p(LC S_{N_d})$ $(p(RC S_{N_d}))$	Relative reinforcement amount, $A_{\beta}$		Expected payoff		
			$A_{\beta L}$	$A_{\beta R}$	EP(L  $S_{N_d}$ )	EP(R  $S_{N_d}$ )	
4	1	1.0000	.833	.167	.833	.000	1.0000
	2	(.0000)	.692	.308	.692	.000	1.0000
	3		.500	.500	.500	.000	1.0000
	4		.308	.692	.308	.000	1.0000
	5		.167	.833	.167	.000	1.0000
3	1	.9125	.833	.167	.763	.014	.982
	2	(.0845)	.692	.308	.634	.026	.961
	3		.500	.500	.468	.042	.915
	4		.308	.692	.282	.059	.828
	5		.167	.833	.153	.071	.683
2	1	.4821	.833	.167	.402	.086	.823
	2	(.5179)	.692	.308	.334	.166	.668
	3		.500	.500	.241	.259	.482
	4		.308	.692	.148	.359	.292
	5		.167	.833	.080	.432	.157
1	1	.1093	.833	.167	.091	.148	.280
	2	(.8907)	.692	.308	.076	.274	.216
	3		.500	.500	.055	.445	.109
	4		.308	.692	.034	.617	.052
	5		.167	.833	.018	.742	.024
0	1	.0256	.833	.167	.021	.162	.116
	2	(.9744)	.692	.308	.018	.300	.056
	3		.500	.500	.013	.487	.026
	4		.308	.692	.008	.675	.012
	5		.167	.833	.004	.812	.005

opportunity and a following choice at all five biasing conditions. The deviation in the left-key peck probability from the average probability,  $\Delta p(L)$ , is shown as a function of the cycle during which the reinforcement opportunity occurred. Here,  $\Delta p(L)$  equals  $p(L|S_n)$  minus the average of  $p(L|S_n)$  with  $D(S_n) = 1$ . In other words,  $\Delta p(L)$  equals  $p(L)$  for a substimulus with a single reinforcement opportunity minus the average  $p(L)$  for all density 1 substimuli. Note that the expected substimulus for density 1 substimuli is (1/4,1/4,1/4,1/4), which is close to substimulus (0,0,0,0). Therefore,  $\Delta p(L)$  reflects not only the deviation of  $p(L)$  from the average  $p(L)$ , but to a greater

extent the effect of adding an opportunity to a given cycle. The value of  $\Delta p(L)$  is largest when a reinforcement opportunity occurs closest to choice, as is the case for substimulus 0001. It rapidly decreases as the opportunity moves away from choice. The change in slope for biases 4 and 5 is less pronounced for Birds 85 and 29, since for these last two cases these birds almost always pecked the right key. For the other birds, a smaller but similar effect is seen for bias 5.

Figure 6 shows  $-\Delta p(L)$  as a function of the distance between a reinforcement omission and a following choice at all five biasing conditions. The value of  $-\Delta p(L)$  equals  $-(p(L|S_n))$

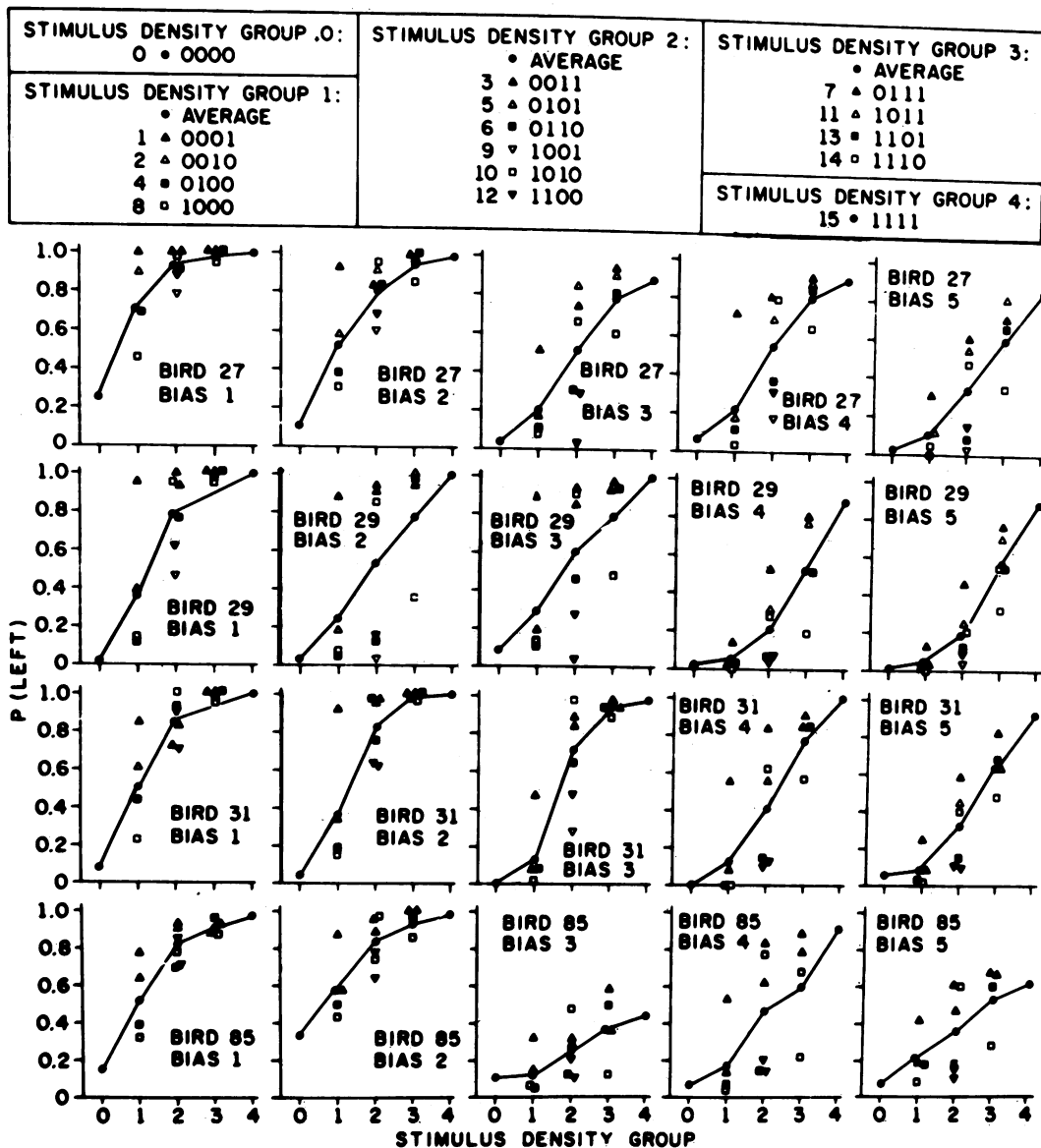


Fig. 4. Left-key-peck probability,  $p(L)$  for each individual substimulus at each biasing, arranged so that all same-reinforcement-density substimuli are grouped above that reinforcement density number. The solid line shows the average  $p(L)$  for each substimulus.

minus the average  $p(L|S_n, D(S_n) = 3)$ ). That is, the right-hand side of the equation is the value of  $p(L)$  for a substimulus with a single reinforcement omission in the  $i$ th cycle, minus the average  $p(L)$  of those density 3 substimuli, all times minus one. This deviation reflects the control exerted by a "missing" reinforcement opportunity. The curves in Figures 5 and 6 have very similar shapes. The biasing effect

in Figure 6 is reversed from that in Figure 5, in the sense that there is more flattening for lower numbered biasing conditions when payoff is greater for a correct left than for a right choice. In either case, the flattening may represent a floor or a ceiling effect at the more extreme values. The birds, ranked in order of increasing sensitivity to reinforcement omission, were Birds 31, 27, 29, and 85, with Birds

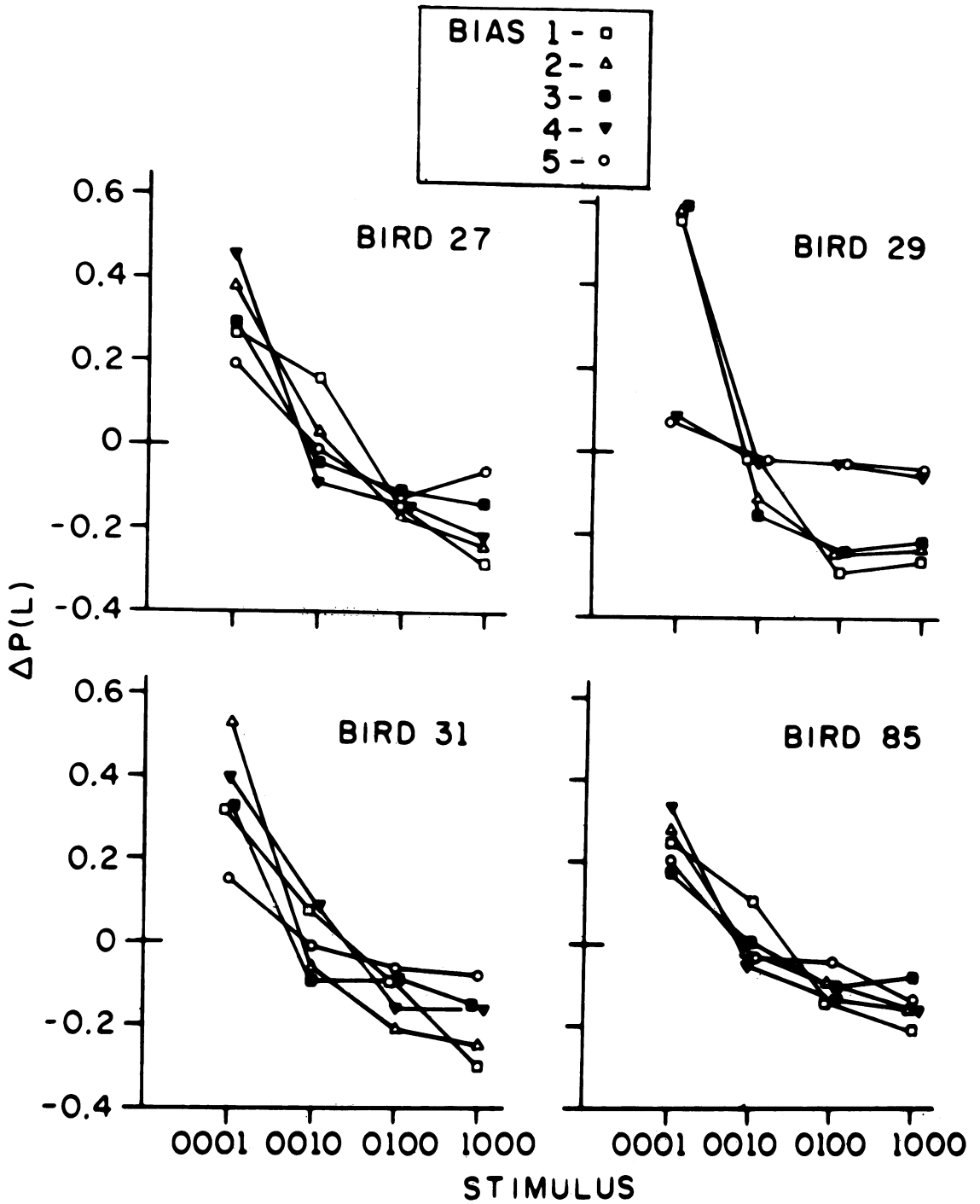


Fig. 5. The positive change in  $p(L)$ , equal to the  $p(L)$  of a substimulus with a single reinforcement opportunity on the  $i$ th cycle, minus the average  $p(L)$  of those density 1 substimuli, plotted as a function of the cycle before choice,  $i$ , in which that reinforcement opportunity fell,  $p(L) = p(L|S_n) - \text{average } p(L|S_n, D(S_n) = N_d = 1)$ . Effect of biasing is also shown.

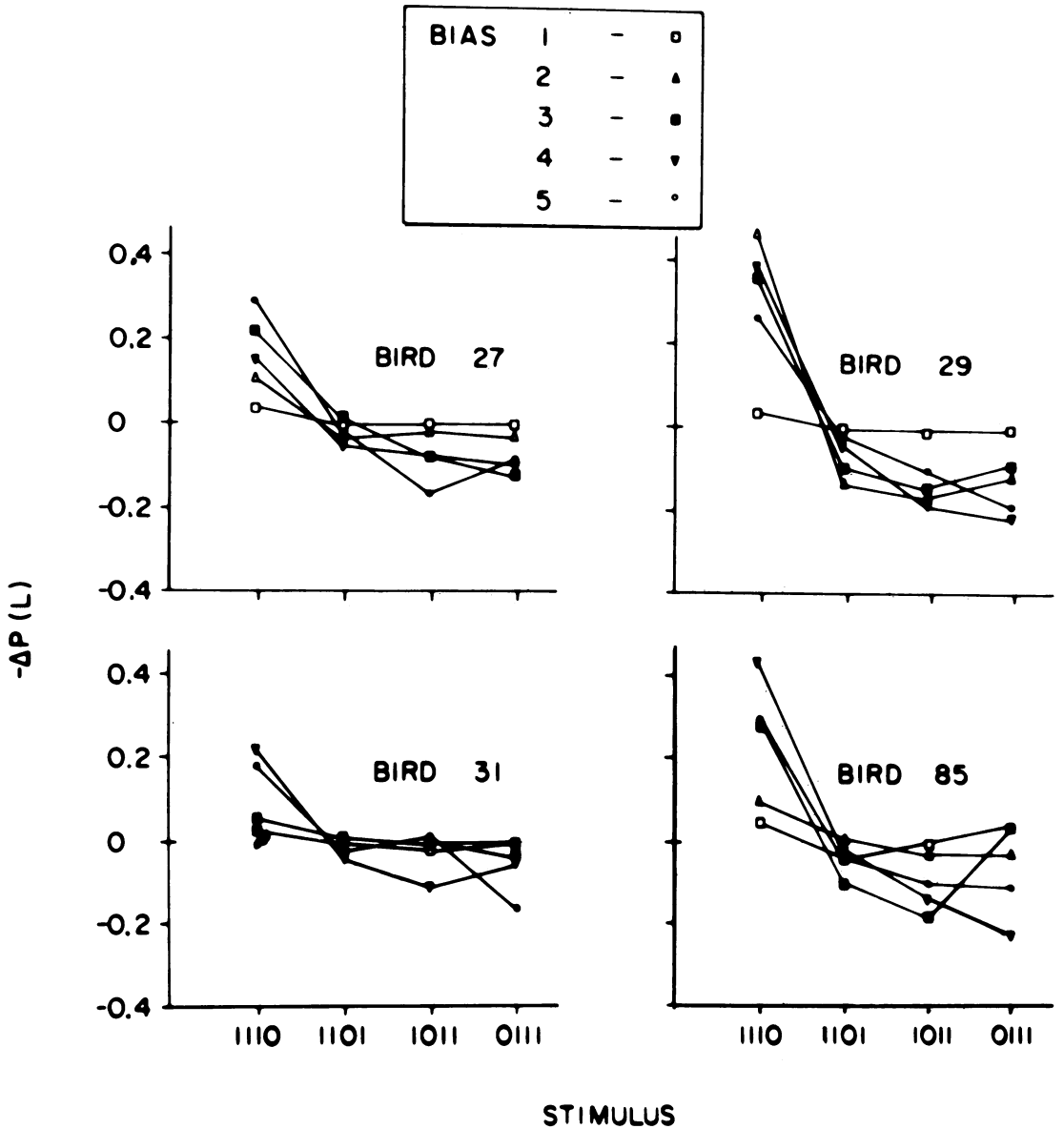


Fig. 6. The negative change in  $p(L)$ , equal to the negative of the  $p(L)$  to a substimulus with a single reinforcement opportunity omission on the  $i$ th cycle, minus the average  $p(L)$  of those density 3 substimuli, is plotted as a function of the cycle before choice,  $i$ , in which that reinforcement opportunity fell (absence case),  $-\Delta p(L) = -(p(L|S_n) - \text{average } p(L|S_n, D(S_n) = N_s = 3))$ . This is the complementary graph to the previous one. The effect of biasing is shown.

29 and 85 being very close. This ranking correlates well with the overall sensitivity as seen in isosensitivity curves in Figure 2.

DISCUSSION

Figure 2 suggests that the sensitivity displayed by three birds was not greatly inferior,

at a gross molar level, to that of an ideal observer. Let us now ask how choice behavior was controlled at the gross molar, molar, molecular, and micro levels, and in what sense birds did not discriminate reinforcement density with maximum possible sensitivity.

The present data suggest that choice behavior was controlled by reinforcement factors,

including both the amount of reinforcement for a correct choice and the number of reinforcement opportunities during the stimulus period, as well as the length of time by which stimulus events preceded choice.

It will be shown here that the decision rules at the molar level for making a left-key peck as a function of the number of reinforcement opportunities in a substimulus can be described as matching-to-relative-expected payoff. This deviation from an optimizing ideal observer's rule in which the response alternative with the higher expected payoff is maximized appears to be due to diminution in the control over choice by earlier events within the substimulus, i.e., by forgetting. This molar matching law fails to account for systematic differences in  $p(L)$  at the molecular level.

Maximizing and matching models will next be contrasted with actual performance at the four different levels of analysis. These contrasts will help to identify a bird's decision rule and to show how it differs from an ideal observer's rule.

A general form of matching and maximizing derived from Herrnstein's (1970) account will be applied at each level. Assume first that, in a discrete-trial choice situation, obtained response probability may be treated as equivalent to the relative rate of responding in a free operant situation. Obtained response probability,  $p(L)$ , may be conditionalized upon a given stimulus condition,  $S_i$ , or  $p(L|S_i)$ . As described above, this stimulus condition is specified differently at each level of description. Assume second that programmed relative reinforcement proportion is equivalent to programmed relative reinforcement rate, rather than obtained relative reinforcement rate. The specification of  $p(L|S_i)$  calls for a generalization of the notion of programmed relative reinforcement proportion to the programmed relative expected payoff for a left-key peck in a given stimulus condition. The relative  $EP(L|S_i)$  is the ratio of the expected payoff for a left-key peck,  $EP(L|S_i)$  to the sum of the expected payoffs for both key pecks:

$$\text{Rel } EP(L|S_i) = \frac{EP(L|S_i)}{EP(L|S_i) + EP(R|S_i)} \quad (1)$$

The expected payoff for a left-key peck in a given stimulus condition,  $EP(L|S_i)$ , may be interpreted as the product of the conditional

probability of reinforcement for a left-key peck in the given stimulus condition,  $p(LC|S_i)$ , and its relative reinforcement amount,  $A_{\beta L}$ :

$$EP(L|S_i) = p(LC|S_i) \cdot A_{\beta L}; \quad (2)$$

$$\begin{aligned} EP(R|S_i) &= p(RC|S_i) \cdot A_{\beta R} \\ &= [1 - p(LC|S_i)] \cdot (1 - A_{\beta L}). \end{aligned} \quad (3)$$

The relative reinforcement amount,  $A_{\beta L} = A_{LC}/(A_{LC} + A_{RC})$ , where  $A_{LC}$  is the amount of reinforcement for a correct left-key peck and  $A_{RC}$  is the amount of reinforcement for a correct right-key peck. Relative expected payoff retains the common notion of programmed relative amount of reinforcement (Catania, 1963).

A decision rule,  $p(L|S_i)^*$ , is designed to predict the obtained probability of a left peck,  $p(L|S_i)$ . Specifically, such a rule predicts the left-key peck proportion under a given set of stimulus conditions as a function of the relative expected payoff for making that response under those stimulus conditions. Two such decision rules will be considered here. One is a maximizing and the other is a matching rule.

To compactly write the maximizing rule, an indicator notion will be introduced (Loève, 1963). The indicator of set  $A$  is the function which assigns the value 1 to all points in  $A$  and value 0 to all points in the complement,  $A'$ . When  $A = \{x|x > k\}$ , the indicator of  $A$  will be denoted by  $I_k(x)$  (sometimes written  $\frac{1}{k}(x)$ ). Here,  $k = .5$  and  $x = \text{Rel } EP(L|S_i)$ . Therefore,

$$\begin{aligned} I_{.5}(x) &= 1 \text{ if } x > .5 \\ &= 0 \text{ if } x \leq .5. \end{aligned}$$

The maximization relation states that  $p(L|S_i)$  is equal to the indicator of the set of all  $\text{Rel } EP(L|S_i)$  whose values are greater than .5. The maximizing rule then is:

$$\begin{aligned} p(L|S_i)^* &= I_{.5} \text{ Rel } EP(L|S_i) \\ &= \begin{cases} 1 & \text{if } \text{Rel } EP(L|S_i) > .5 \\ 0 & \text{if } \text{Rel } EP(L|S_i) \leq .5, \end{cases} \end{aligned} \quad (4)$$

and the matching rule is:

$$p(L|S_i)^* = \text{Rel } EP(L|S_i). \quad (5)$$

At the gross molar level, two sets of decision rules might apply depending on the extent to which the rich- and lean-schedule stimuli were discriminable from one another. The case where the two stimuli are not discrimi-

nable is equivalent to a probability learning situation with  $p(L|S_i)^* = p(L)^*$  since payoff-controlled response bias completely overshadows probability of stimulus-presentation controlled bias (Nevin, 1969). The corresponding first form of maximizing,  $p(L)^* = I_{.5} \text{ Rel } EP(L)$  is called overall response alternative maximizing where, irrespective of stimulus differences, one alternative is always chosen. This may be contrasted with a first form of matching,  $p(L)^* = \text{Rel } EP(L)$ , which is called overall payoff matching, where an alternative is chosen in proportion to overall payoff, a degenerate form of expected payoff. The case where the stimuli are perfectly discriminable may be idealized by a second form of either maximizing or matching. According to conditional maximizing,  $p(L|S_i)^* = I_{.5} \text{ Rel } EP(L|S_i)$  where  $i = \text{lean, rich}$ , one response alternative is always chosen on trials where a particular stimulus is presented (Coulson, Koffer, & Coulson, 1971; Herrnstein & Loveland, 1975). This is the same as matching the probability of a response to the stimulus presentation probability,  $p(L|S_i)^* = p(S_i)$ , a degenerate form of matching to relative expected payoff when relative reinforcement amount is disregarded.

Figure 3 showed that as biasing increased, the overall bias,  $p(L)$ , linearly increased. This result suggests imperfect discrimination at the gross molar level. The slope of the lines suggests a slight degree of overall payoff undermatching in the direction away from overall response alternative maximizing. Conditional maximizing would not be expected because an ideal observer could only imperfectly discriminate between the two overlapping probabilistic stimuli. Although matching in the probability learning sense at the gross molar level seems more clearly indicated than conditional maximizing for unequal biasing, the birds did discriminate between the two stimuli. It is advisable, therefore, to continue the analysis to see how overall matching arises when there is this discrimination.

Both a maximizing model and a matching model are considered next at the molar level. For an optimizing ideal observer, the maximizing model is:

$$p(L|S_{N_d})^* = I_{.5} \text{ Rel } EP(L|S_{N_d}),$$

$$N_d = 0, 1, 2, 3, 4; S_i = S_{N_d}, \quad (6)$$

and  $p(L|S_{N_d})^*$  is independent of the proximity

of a reinforcement opportunity to a following choice. Sensitivity could be reduced by any deviation from this ideal decision rule due to any combination of control by: (a) proximity of events to choice, (b) biasing, and (c) a ratio of expected payoff for the optimal response relative to the other response too close to one. To maximize reinforcement for choice within the range of biases used here, the optimizing ideal observer always pecks right for density 0 and 1 substimuli, left for density 3 and 4 substimuli, and pecks either right or left for density 2 substimuli, depending on whether biasing was to the right or left. This ideal decision rule maximizes reinforcement for all biasing since  $\text{Rel } EP(L|S_{N_d}) < .3$  and  $\text{Rel } EP(R|S_{N_d}) > .7$  for density 0 and 1 substimuli, and  $\text{Rel } EP(R|S_{N_d}) < .3$  and  $\text{Rel } EP(L|S_{N_d}) > .7$  to density 3 and 4 substimuli, as was shown in Table 2. For density 2 substimuli, an ideal observer always maximizes by pecking the higher payoff side key. This conditional maximizing could produce matching of the response probability to stimulus presentation probability, as described previously.

At the molar level, empirical evidence of Terman and Terman (1972), Coulson et al. (1971), and Herrnstein and Loveland (1975) suggests that an alternative will be maximized in a noncorrection situation as long as relative reinforcement rate for alternative responses differs significantly from .5. Also, the same key should be pecked following events with the same value, where value is measured along a critical dimension such as reinforcement density. If substimulus reinforcement density alone controls choice, one response alternative would be maximized.

A value of  $\text{Rel } EP(R_i|S_i)$  somewhat close to .5 does not always interfere with maximizing. For instance, rats maximized  $p(L)$  when  $\text{Rel } EP(L) = .6$  in a situation where a left-bar press led to a variable-interval delay of 10 sec followed by a differential-reinforcement-of-other-behavior 5-sec schedule (*tand* VID 10 DRO 5), and a right-bar press led to a *tand* VID 25 DRO 5-sec schedule (Coulson et al., 1971). In the present case, for low-reinforcement density substimuli,  $p(L)$  should tend toward 0 and for high density substimuli, toward 1, since correct choices were differentially reinforced, and differences in the temporal location of reinforcement for a substimulus were irrelevant,

as were the variations in the comparison stimuli in the Coulson et al. experiment.

Pigeons on a concurrent variable-ratio variable-ratio schedule maximized reinforcements per response by pecking the key correlated with the smaller ratio so that  $p(L)$  often approximated 0 or 1 (Herrnstein & Loveland, 1975). In the present study, both alternative key pecks were often reinforced for almost every substimulus and for same-density substimuli since the rich and lean schedule overlapped. However, this should not have much effect on maximization as long as relative amount of reinforcement for left- versus right-pecks is not too close to .5 (Herrnstein & Loveland).

The points connected by the solid lines in Figure 4 illustrate that the empirical decision rules deviated from the ideal observer's maximizing decision rule, Equation 6, in three ways. First, although density 0 and 4 substimuli produced correct choices almost always by minimizing  $p(L)$  in the former and maximizing  $p(L)$  in the latter case, density 1 and 3 substimuli, which should have produced the same  $p(L)$  as 0 and 4, respectively, did not. Hence, at the molar level density was not discriminated as an ideal observer would. Second, the proximity of a reinforcement opportunity to choice produced a large ordered deviation in each  $p(L|S_n)$  from the average  $p(L|S_n)$  of same-density substimuli, as shown in Figures 4, 5, and 6. Third, the probability of pecking a given key to same-density substimuli was neither the same nor maximized for all same-density substimuli (except when there were no other substimuli of the same density). This indicated that the birds neither "counted" reinforcement opportunities nor perfectly discriminated density; either each substimulus was not discriminated or one response alternative for a given substimulus was not maximized.

Instead of behaving like ideal observers, the birds behaved as if they had a second strategy which minimized the reinforcement loss for choice that is produced by poor memory for events in the stimuli. With this second strategy, they matched some central tendency of  $p(L)$ , e.g., the average  $p(L)$  or the median  $p(L)$  for same density substimuli, to the relative expected payoff for those same density substimuli. The matching model's predicted left-peck probability,  $p(L|S_i)^*$ , equals the relative expected payoff for a left-key peck to a given density of reinforcement substimulus. The  $S_i$

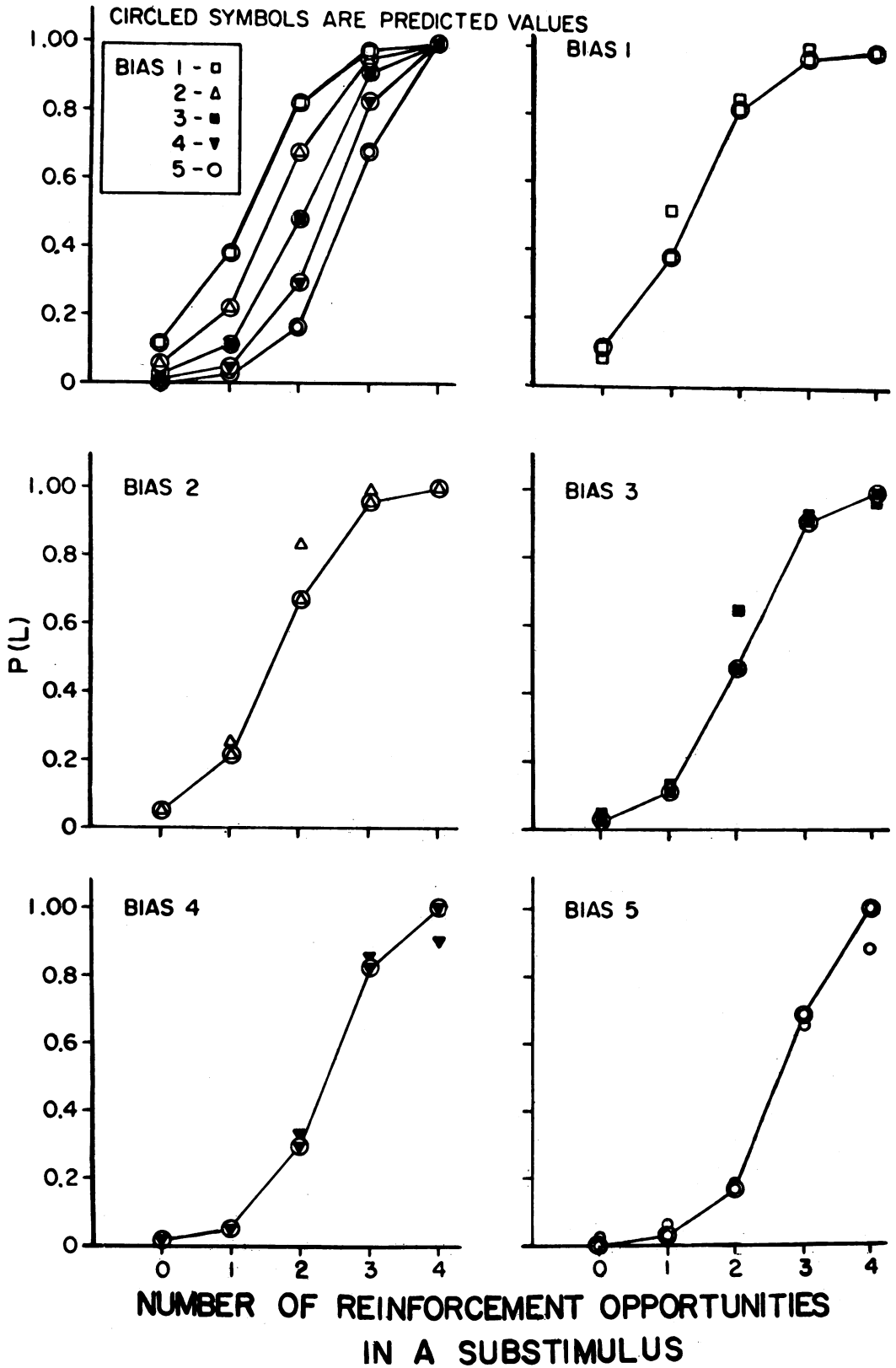
in this case are all the substimuli with the same density,  $S_{N_d}$ .

$$\begin{aligned} p(L|S_{N_d})^* &= \text{Rel } EP(L|S_{N_d}), \\ D(S_n) &= N_d, \\ N_d &= 0, 1, 2, 3, 4. \end{aligned} \quad (7)$$

The predicted  $p(L|S_{N_d})^*$  versus substimulus reinforcement density is plotted in the top left panel of Figure 7. The biasing parameter generates the five curves, each plotted separately in its own panel.

This prediction in effect states that a subject matches its choice to the relative expected values of the alternatives (Herrnstein, 1970). If a correction procedure were used, instead of the present noncorrection procedure, so that an incorrect choice did not "waste" a reinforcement opportunity, the present model would reduce to Herrnstein's model. Then the probability of a trial ending with a correct peck, given any substimulus, would be one, and only relative bias would have an effect. The biasing does not change the shape of the theoretical  $p(L|S_{N_d})^*$  decision curves, it just shifts them. Also, sensitivity remains constant across changes in biasing. Temporal location of reinforcement opportunities within a substimulus should have no effect.

On a molar level, consider that the left-key peck probability for each same-density substimulus,  $p(L|S_{N_d})$ , closely matches the relative expected payoff for left-key pecks to those same-density substimuli. Compare the average  $p(L|S_{N_d})$  lines in Figure 4 to the predicted  $p(L|S_{N_d})^*$  lines in Figure 7. Since the relationship between  $p(L|S_n)$  and reinforcement opportunity proximity is nonlinear, as was shown in Figures 5 and 6, the median  $p(L|S_n)$  at each biasing value and each substimulus density better describes the decision rules than the average shown in Figure 4. However, the values obtained with the two measures did not differ very much. A measure of central tendency that represents performance at the molar level masks the effect of individual same-density substimuli. In Figure 7, the median  $p(L|S_n)$ ,  $D(S_n) = N_d$  for Birds 27, 29, and 31, is compared to the expected  $p(L|S_{N_d})$  for each bias. Bird 85's data were omitted for three reasons: (a) the effect of biasing was disorderly as shown by the anomalous point in Figure 3, and in the bottom panel, biasing condition 3,





in Figure 4; (d) there was not much of an effect of including Bird 85's data other than at the anomalous point; and (c) actual data points were preferred as representatives of central tendency, rather than the average of two values. This situation is possible only when using three birds. The predicted points are the circled points connected by lines. Where the obtained points fall on the predicted values they are omitted. Only five (20%) of the obtained points differed by more than 5% from the predicted  $p(L|S_N)_d^*$  values.

The overall matching obtained at both gross molar and molar levels was an average produced by more left pecks to substimuli with reinforcement closer to choice than to substimuli having the same reinforcement but with reinforcement further from choice.

Matching at the molar level should be the second best strategy. Shifts away from overall matching produce lower total amounts of reinforcement. If a bird shifts to the left more for 1110 to maximize reinforcement for choice given that substimulus, then it also shifts to the left for 0001. What is gained by the shift for 1110 is more than offset by the loss for the shift at 0001. This assumes that a bird discriminates density as well as possible given a limited memory. To have constant overall sensitivity, changes in response probability to one substimulus must result in changes in response probability to all. A bird cannot shift bias for just one or two substimuli because that would change its overall sensitivity to reinforcement density. It can only shift overall bias. This forces a bird to set the overall bias and establish an anchor point. The  $Rel EP(L|S_N)_d$  sets the bias for a set of same-reinforcement density substimuli, thereby setting the overall bias.

This rationale for matching does not, however, explain how matching behavior is generated. Local maximizing in some form would occur if the criteria for choice were conditional on, for instance, certainty about substimuli. There is no way to distinguish local maximizing of reinforcement from molecular matching in the present study because there is no independent estimate of the perceived value of a substimulus; nor is there an inde-

pendent estimate of the degree of impaired sensitivity to the events within a substimulus. One property of substimuli is known, however—the temporal property.

The effect of concatenating events in a temporal sequence must be taken into account to explain  $p(L|S_n)$  at the molecular level. First, the temporal properties of  $S_n$  will be discussed. At this level temporal factors partially account for the deterioration of reinforcement-density discrimination. For same-density substimuli,  $p(L)$  decreased as the reinforcement opportunities occurred further from choice, as was seen in Figure 5. For the obverse case,  $p(L)$  increased to a smaller extent as the missing reinforcement opportunity occurred further from choice, as was seen in Figure 6. Reinforcement opportunity occurrence exerted more control over choice than nonoccurrence.

At the micro level, control by a single event preceding choice decreased as a function of the event's temporal separation from choice. In Figure 8, control by an event  $x$  seconds before a choice in the present experiment is compared to control in a delayed matching-to-sample experiment by Berryman, Cumming, and Nevin (1963). There, the sample was either a red, green, or blue center key. An observing center-key peck turned it off and, after a delay, turned on the two side-key comparison stimuli. A peck on a side key with a hue that matched the center key's hue produced reinforcement; a nonmatching side-key peck was followed by a blackout.

To directly compare these results, comparable strength-of-control measures must be used. In the Berryman et al. (1963) experiment, the correct choice probability,  $p(C)$ , may be said to define stimulus control by a sample event. As time between the sample event and the choice increased from 0 to 25 seconds, the median  $p(C)$  for three birds decreased from approximately 1.0 to .5.

Matters were more complicated in the present study. In the unbiased condition, the event whose control over choice was examined was either a single addition of a reinforcement opportunity on one of four cycles (density 1 substimuli) or a single omission of a reinforcement opportunity on one of four cycles (den-

Fig. 7. Median left key peck probability for three birds, median  $p(L|S_N)_d$ , to same density substimuli. The curves shown in panel 1 are the expected left-key-peck probabilities,  $p(L|S_N)_d^*$ , given Equation 7.

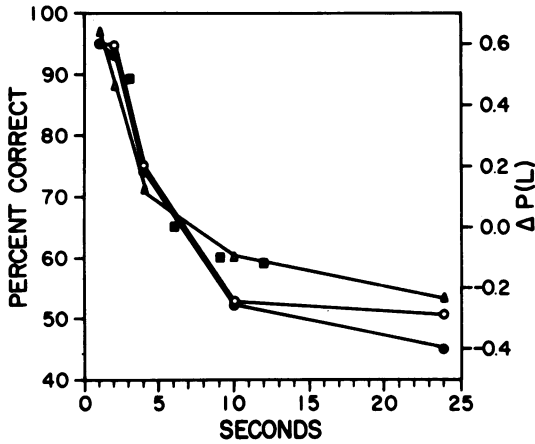


Fig. 8. The left y-axis has the median probability of a correct match-to-sample for three pigeons. The triangles represent that median plotted against different delays between the discriminative stimulus and the choice, from Berryman, Cumming, and Nevin (1963). The right-hand axis has the probability of a correct discrimination scaled in the reverse manner. The independent variable in the second case is time before choice of the singular occurrence of a reinforcement opportunity for stimuli with a reinforcement density of 1, and time before choice for the singular non-occurrence of a reinforcement opportunity for stimuli with a reinforcement density of 3.

sity 3 stimuli). As time between either the single addition or omission and choice increases,  $p(C) = p(LC) + P(RC)$  increases.

These increases in control show up as increases in the probability of an error,  $p(E)$ . For density 1 stimuli, an ideal observer would always peck the right key, since  $EP(L|S_n) < .055$ ; any increase in  $p(L)$  as the single event occurred closer to choice, would lead to slight increases in  $p(LC)$  and to large decreases in  $p(RC)$  with their sum,  $p(C)$ , decreasing. Similarly, as the omitted reinforcement opportunity occurred closer to choice,  $p(L)$  would decrease and error probability,  $p(E)$ , would increase. Hence, an increase in control by these single events would be reflected in a decrease in  $p(C)$ , which would go from approximately .9 for no control to .5 for the maximum control achieved in the present experiment. The maximum control would occur for the shortest time span; that is, the cycle only 3 sec away from choice, with  $p(C) = .5$ . To read strength of control directly in both experiments, the right hand y-axis of Figure 8 was inverted. The median strength of single event control for Birds 29, 31, and 85 in the present experiment (on the right hand y-axis) may then be

directly compared to the median strength of discriminative control for Birds 170, 171, and 172 in the Berryman et al. (1963) experiment (on the left-hand y-axis), as a function of the intervening interval size. The three curves have similar shapes, the upward displacement of the extra reinforcement curves from the missing reinforcement curves being due to the differential effect of an additional event on a blank background versus a missing event on a background involving reinforcement.

Two models show how results from the molar and micro levels may be combined to explain performance at the molecular level. Because the class of same-density substimuli has only one member each for density 0 and 4, any predictions at the molar level for them are identical to predictions at the molecular level; therefore, differences between those predictions are trivial. For density 1, 2, and 3, there are four, six, and four substimuli, respectively. In the first model, a weighted mean value of  $p(L|S_n) = N_a$  is set by the matching law, and the deviations from this weighted mean are set by some forgetting law. For density 1 and 3 substimuli the memory curves are relatively simple. Both the curve for single reinforcement events (density 1) and the flatter curve for single missing reinforcement events (density 3) show the decrement in control exerted by events over time as found in delayed-matching-to-sample studies. For density 2 substimuli, there are two alternative approaches. Either a third kind of forgetting curve has to be invoked to describe control by the compound of two reinforcement opportunities and two missing reinforcement opportunities, or control of the ensemble of events has to be predicted from how each event's control over choice combines with the control by other events. A memory scale can be constructed from the concatenation of such events (Krantz, Luce, Suppes, & Tversky, 1971). Alternatively, in a second model, the remembered value of each substimulus is probabilistically determined, with each value of  $S_n$  being a random variable with a mean and a variance. One response alternative would be maximized as a function of the perceived relative expected payoff which changes trial to trial, an idea similar in some respects to Shimp's (1976) momentary maximizing model, the average across trials equaling that predicted by the matching law.

## APPENDIX

Two models of an ideal observer could be assumed: one based on a discrete binomial distribution, and the other based on a pointwise discontinuous binomial distribution. The discrete form has probability represented by a vertical line at each possible integral substimulus density value,  $N_d$ , that could serve as a criterion. The pointwise discontinuous (almost-everywhere continuous) distribution has probability represented by bars centered over these same criterion points, extending  $\frac{1}{2}$  a reinforcement opportunity to the left and to the right. If the discrete distribution is assumed, each isosensitivity function consists of a set of discrete points as shown by the  $x$ 's on the top left isosensitivity curve,  $d' = 2/\sigma$ . To find P(Hit) and P(False Alarm) for a given criterion,  $N_d$ , and separation between means,  $\Delta M$ , the discrete binomial probability function is integrated from the criterion value to plus infinity. For P(Hit), the discrete probability function for the rich stimulus is integrated by summing all values of P(Hit) at substimulus density values,  $N_d$ , above the criterion. The value of P(Hit) at a given substimulus density value,  $N_d$ , is equal to the frequency of  $LC$  divided by the frequency of  $LC + RE$  at that density. For P(False Alarm) the discrete probability function for the lean stimulus is integrated by summing each value of P(False Alarm) determined at a given  $N_d$  value above the same criterion.

A second ideal observer model can be represented as having only a pointwise discontinuous probability density distribution. In the only pointwise discontinuous case, the discrete binomial distributions are modified in order to obtain isosensitivity points intermediate to values arising at the discrete criterion points that were obtained from the discrete binomial probability distributions. To connect these isosensitivity points, probability at a given integral value of  $N_d$  in the probability distribution is replaced by the numerically proportional probability density. The probability density at the left of an interval,  $p(S_{N_d})$ , is assigned to non-integral values within the interval,  $N_d$  to  $N_{d+1}$ . This converts the discrete probability distribution into a probability density distribution (as was done by Egan, 1975). To find the theoretical values of P(Hit) and P(False Alarm) for the only pointwise dis-

continuous case, one integrates over each continuous portion up to the criterion and then sums the integrals of these portions.

Different  $d'$  values were assigned to the discrete and the only pointwise discontinuous isosensitivity functions. First,  $d'$  was calculated in the discrete case. For the two functions lower on the graph, the two additional  $d'$ 's were then found by either assuming a fixed standard deviation while the difference between means of the two distributions,  $\Delta M$ , gets smaller as sensitivity deteriorates, or by assuming that the degradation in sensitivity is due to a decrease in the difference between  $p$  and  $q$ , where  $q = 1 - p$ .

Three theoretical sets of isosensitivity points in the discrete probability case, corresponding to  $d' = 2.31, 1.73$ , and  $1.15$ , were obtained in the following fashion. The value of  $d' = (M_1 - M_2)/\sigma$  (Egan, 1975), were found for the distributions when their means were maximally separated (2.0 reinforcement opportunities); moved together by  $\frac{3}{4}$  maximal separation (1.5); and by  $\frac{1}{2}$  maximal separation (1.0). The means,  $M_{s_{rich}}$  and  $M_{s_{lean}}$ , or  $M_1$  and  $M_2$ , were found from  $M = Np$ , where  $M = N_d$ , the density of the gross level stimulus,  $N =$  number of cycles, and  $p =$  probability of reinforcement for the first center-key peck in a cycle. For the maximum separation of the two distributions,  $\Delta M = M_{s_{rich}} - M_{s_{lean}}$ , is calculated as follows:  $M_{s_{rich}} = 4 \cdot \frac{3}{4} = 3 = N_d$  for the rich density substimuli;  $M_{s_{lean}} = 4 \cdot \frac{1}{4} = 1 = N_d$  for the lean density substimuli, and  $\Delta M = M_{s_{rich}} - M_{s_{lean}} = 3 - 1 = 2.0$ . For three sets of points, a fixed standard deviation,  $\sigma = \sqrt{Np(1-p)} = \sqrt{4(\frac{3}{4})(\frac{1}{4})} = \sqrt{\frac{3}{2}} = .866$  was assumed for this first discrete case. Since the inverse,  $1/\sigma = 1.1547$ , for the constant  $\sigma$ , and  $d' = 1/\sigma \cdot \Delta M$ ,  $d'$  values were 2.31, 1.73, and 1.15, when  $\Delta M = 2.0, 1.5$ , and  $1.0$ , respectively.

In the only pointwise discontinuous case, the area from the criterion to infinity is the same as the corresponding sum in the discrete case so that P(Hit) and P(False Alarm) are the same at the criterion points and are linear in between. The  $d'$  values assigned to the isosensitivity curves shown in Figure 2 are derived from the only pointwise discontinuous probability distribution case next. These  $d'$  values in the pointwise discontinuous case are

2.19, 1.64, and 1.10, which are smaller than the corresponding isosensitivity point values. Remember that  $d' = 1/\sigma \cdot \Delta M$ . While the difference between the means is the same in the discrete and only pointwise discontinuous case, the standard deviation in the former,  $\sigma = .866$  is smaller than the  $\sigma = 1.10$  of the latter, as will be shown for the only-pointwise discontinuous case. The mean,  $\mu$ , used to calculate the variance and the corresponding standard deviation is found from:

$$\mu = \sum_{n=0}^4 \int_n^{n+1} xp(x)dx = 3.5,$$

for the rich stimulus and 1.5 for the lean stimulus. Note that the maximal separation,  $\Delta M = 2 = 3.5 - 1.5$ , which was the same as in the discrete case.

The variance is found from:

$$\sigma^2 = \sum_{n=0}^4 \int_n^{n+1} (x - \mu)p(x)dx.$$

Instead of setting a fixed  $\sigma$ , if the degradation in sensitivity were due to moving  $p$  towards  $q$ ,  $\Delta M$  values would be the same. As  $p - q$  approaches 0, in the discrete case  $\sigma$  approaches 1, which is not very different from .866, and in the almost continuous case,  $\sigma$  would approach 1.04, which is not very different from 1.10. The values of  $d'$  in either case would not be very different.

## REFERENCES

- Berryman, R., Cumming, W. W., & Nevin, J. A. Acquisition of delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior*, 1963, 6 (1), 101-107.
- Catania, A. C. Concurrent performances: a baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior*, 1963, 6, 299-301.
- Catania, A. C., & Reynolds, G. S. A quantitative analysis of responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 327-383.
- Coulson, G., Koffer, K., & Coulson, V. Maximization of pellet density by rats in a two-choice operant situation. *Psychonomic Science*, 1971, 24(3), 145-146.
- Egan, J. P. *Signal Detection theory and ROC analysis*. New York: Academic Press, 1975.
- Estes, W. K., Burke, C. J., Atkinson, R. C., & Frankmann, J. P. Probabilistic discrimination learning. *Journal of Experimental Psychology*, 1957, 54(4), 233-239.
- Herrnstein, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, 13, 243-266.
- Herrnstein, R. J., & Loveland, D. H. Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior*, 1975, 24, 107-116.
- Hobson, S. L. Discriminability of fixed-ratio schedules for pigeons: effect of absolute ratio size. *Journal of the Experimental Analysis of Behavior*, 1975, 23, 25-35.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. *Foundations of Measurement*. (Vol. 1) New York: Academic Press, 1971.
- Lattal, K. A. Reinforcement contingencies as discriminative stimuli. *Journal of the Experimental Analysis of Behavior*, 1975, 23, 241-246.
- Loève, M. *Probability Theory*. (3rd Ed.) Princeton: Van Nostrand, 1963.
- Nevin, J. A. Signal detection theory and operant behavior. Review of D. M. Green and J. A. Swets *Signal detection theory and psychophysics*. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 475-480.
- Pliskoff, S. S., & Goldiamond, I. Some discrimination properties of fixed-ratio performance in the pigeon. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 1-9.
- Rilling, M., & MacDiarmid, C. Signal detection in fixed-ratio schedules. *Science*, 1965, 148, 526-527.
- Schoenfeld, W. N., Cumming, W. W., & Hearst, E. On the classification of reinforcement schedules. *Proceedings of the National Academy of Sciences*, 1956, 42, 563-570.
- Shimp, C. P. Probabilistic discrimination learning in the pigeon. *Journal of Experimental Psychology*, 1973, 97, 292-304.
- Shimp, C. P. Short-term memory in the pigeon: relative recency. *Journal of the Experimental Analysis of Behavior*, 1976, 25, 55-62.
- Swets, J. A., Tanner, W. P., & Birdsall, T. G. Decision processes in perception. In J. A. Swets (ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964.
- Terman, M., & Terman, J. S. Concurrent variation of response bias and sensitivity in an operant psychophysical test. *Perception and Psychophysics*, 1972, 11(6), 428-432.
- Weissman, A. Impairment of performance when a discriminative stimulus is correlated with a reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 1961, 4, 365-369.

Received November 8, 1976

Final acceptance January 30, 1979