OPERANT CONDITIONING OF BEHAVIORAL VARIABILITY USING A PERCENTILE REINFORCEMENT SCHEDULE

ARMANDO MACHADO

UNIVERSITY OF LISBON, PORTUGAL

The present investigation developed and tested a new percentile reinforcement schedule suited to study pattern variability, whose main feature was the relative dissociation it provided between the variability requirement defining criterional responses and overall probability of reinforcement. In a discrete-trials procedure, pigeons produced patterns of four pecks on two response keys. If the pattern emitted on the current trial differed from the N preceding patterns, reinforcement was delivered with probability u. The schedule continuously adjusted the criterion N such that the probability of a criterional response, estimated from the subject's recent behavior, was always constant. In these circumstances, the criterion corresponded to an invariant percentile in the distribution of recent responses.

Using a between-subjects design, Experiment 1 manipulated the variability requirement—the percentile—while keeping overall reinforcement probability constant. The degree of variability varied directly with the requirement. In addition, an inverse relationship existed between the requirement and within-group variance. Experiment 2 manipulated probability of reinforcement while maintaining the variability requirement constant. No consistent relationship was found between variability and reinforcement probability. A tentative hypothesis was advanced ascribing the operant conditioning of behavioral variability to a process of *probability-dependent selection*.

Key words: operant variability, intermittent reinforcement, percentile reinforcement schedule, response patterns, key peck, pigeons

The usual procedure to condition behavioral variability involves the delivery of a reinforcer whenever the current response differs from one or more previously emitted responses (e.g., Morris, 1987; Neuringer, 1986; Page & Neuringer, 1985; Pryor, Haag, & O'Reilly, 1969; Schoenfeld, Harris, & Farmer, 1966). Results from these studies suggest that variability is a property of behavior sensitive to its consequences and, therefore, amenable to operant conditioning. The differential nature of reinforcement necessary for shaping, however, implies intermittency. Because nondifferential but intermittent reinforcement has been shown to increase variability (see Boulanger, Ingebos, Lahak, Machado, & Richelle, 1987; Galbicka,

1988, for reviews), changes in variability during a differentiation can as easily be ascribed to intermittency per se as to shaping whenever reinforcement frequency is allowed to vary. Therefore, a clear demonstration of the operant conditioning of behavioral variability is possible only when the effects of reinforcement intermittency are controlled adequately. Page and Neuringer (1985; Experiment 5) are the only authors that have addressed this problem, using a self-yoked control design. Initially, trials consisted of patterns of eight pecks on either of two response keys. If the current pattern differed from the patterns emitted during the previous 50 trials, food was delivered. This procedure engendered a high degree of behavioral variability. Next, the sequence of reinforced and unreinforced trials obtained under the last six sessions of this procedure was repeated without any variability requirement (i.e., variability was permitted but not demanded). Behavioral variability decreased sharply, notwithstanding the same intermittency of reinforcement. It was concluded that the initially observed variability was not a by-product of the schedule of reinforcement.

When the variability requirement is manipulated (e.g., Page & Neuringer, 1985; Experiment 3) such that the current pattern must be novel relative to the patterns during the preceding N trials and where N varies across

phases, the yoked procedure is of limited value in controlling the changes in reinforcement probability. Suppose two different criteria (N_1 and N_2) generate different amounts of response variability when individually programmed, but also different probabilities of reinforcement (r_1 and r_2). The associated yoked conditions might show that r_1 and r_2 alone, with no requirement to vary the response patterns, maintain smaller amounts of variability. From this evidence, one could conclude that N_1 and N_2 were both needed to sustain higher degrees of response variation, but it would be difficult, if not impossible, to assess the differential effects of r. An intimately related issue is the possible interaction between N and r whereby the effects of one variable, say r, would be dependent on the other, N. In short, the yoked procedure can certainly demonstrate a global effect of N as in Page and Neuringer's Experiment 5, but it is incapable of controlling differential effects of r across variations of N. The proper control procedure must allow the manipulation of the variability criterion, N, while keeping probability of reinforcement constant. Conversely, to study the interaction between explicitly shaped variability and that arising as a by-product of intermittent reinforcement, the probability of reinforcement must be experimentally manipulable while the variability requirement is kept constant, the opposite of the first point. In the procedure described above, probability of reinforcement cannot be an independent variable because it is partially determined by the subject's behavior.

Developing and testing a new procedure that allows independent control over probability of reinforcement and variability requirement were the main goals of the present research. Probability of reinforcement can be handled adequately if the experimenter controls the probability of a criterional response. However, this is not possible if the criterion is specified in terms of a physical constant (e.g., in terms of a fixed number of trials without a given sequence) because learning continuously changes the proportion of the subject's behavior meeting the criterion. To elucidate the difficulty consider the situation described earlier that constituted the starting point of this research: In a discrete-trials procedure, pigeons are required to generate patterns of, say, four pecks on two response keys. The pattern produced on trial N is then compared with the patterns

produced on trials $N - 1$, $N - 2$, and so on, until a match is found. The number of intervening trials is the variability score of the current pattern. Another way to think about this task is to consider that only 16 sequences are possible (2^4). Consequently, the current sequence has most likely occurred in the past, and the number of trials between its last occurrence and the present recurrence is the variability score of the current sequence. Assuming no sequential dependencies from trial to trial (see Bryant & Church, 1974; Page & Neuringer, 1985), the probability of a particular variability score x is given by the geometric distribution

$$P(x) = p(1 - p)^x, \qquad (1)$$

where p stands for the probability of a repetition (see Figure 1). Greater variability means a smaller p and, when all sequences are equally likely, p is $1/16$, .0625. If variability scores must exceed a fixed value prior to reinforcement, as in Page and Neuringer's experiments, then increases in variability will increase the proportion of the subject's behavior meeting the criterion, concomitantly increasing the overall reinforcement probability (provided all criterional responses are reinforced). This problem is critical if the effects of different criteria are being compared because criterion changes will be interwoven with the effects due to concurrent changes in reinforcement probability.

One alternative that allows control of reinforcement probability is to adjust the criterion relative to the current level of variability being exhibited. In other words, criterional variability scores may be increased or decreased, as behavior becomes more or less variable, such that the expected probability of a criterional response (i.e., the probability the next variability score will exceed the criterion) is held constant. To accomplish this, the distribution of the subject's recent variability scores is used as a sample from which an experimentally specified *percentile* point is calculated. This percentile is then used as the variability requirement on the next trial. By continuously updating the sample—dropping the oldest and adding the most recent variability score—and adjusting the criterion as specified, the procedure controls the probability of a criterional response. This probability is the complement of the percentile (e.g., to

obtain criterional responses with probability .3, the 70th percentile is specified). Note that although the cardinal value of the criterion changes as differentiation proceeds, its ordinal value, relative to the subject's current distribution of responses, remains constant because the criterion always demarcates the same percentile value. In order to have different variability requirements the experimenter specifies different percentile values. To obtain a fixed probability of reinforcement per trial, $P(S^+)$, the experimenter adjusts the conditional probability of reinforcement given a criterional response, $P(S^+|Cr)$, to the unconditional probability of a criterional response, $P(Cr)$, according to the equation

$$P(S^+) = P(Cr)P(S^+|Cr). \qquad (2)$$

(Note that noncriterional responses were never reinforced.) Thus, to obtain an overall probability of reinforcement, $P(S^+)$, of .3 having chosen the 40th percentile (i.e., $P(Cr) = .60$) the $P(S^+|Cr)$ would be set at .50.

The schedule so defined corresponds to a percentile reinforcement schedule (see Davis & Platt, 1983; Galbicka & Platt, 1989; Scott & Platt, 1985; and especially Galbicka, 1988, in press; and Platt, 1973, for a thorough discussion of percentile reinforcement schedules). Finally, as sample percentile points estimate the corresponding percentile points of the population from which the sample was drawn, percentile schedules control the probability of a criterional response irrespective of the shape the response distribution might have.

The remainder of this article presents two experiments in which the effects of the variability requirement and probability of reinforcement were dissociated and independently assessed.

EXPERIMENT 1

Page and Neuringer (1985, Experiments 3 and 5) not only found that variability increased when patterns unlike those produced during the *N* preceding trials were differentially reinforced, but also that the degree of variability directly varied with *N*. The present experiment attempted to reproduce this relationship between required and obtained variability using a percentile reinforcement schedule that maintains constant reinforcement probability. To attain this goal, four groups of birds were
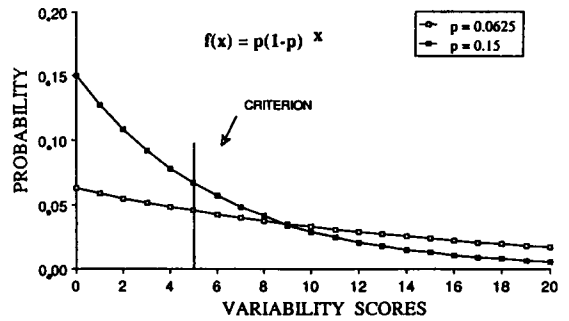

Fig. 1. Probability distribution of variability scores when performance is random (probability of a repetition $p = .0625$) or stereotyped ($p = .15$, for example). Scores exceeding the criterion are followed by reinforcement.

subjected to decreasing variability requirements but to the same overall probability of reinforcement. If variability is a differentiable response property it was expected that Groups 1 to 4 would be ranked in this order on the variability measures.

METHOD

Subjects

The subjects were experimentally naive homing pigeons (*Columba livia*). Each pigeon was maintained at 80% (±10 g) of its free-feeding body weight. Water and grit were available continuously in their home cages and a 12:12 hr light/dark cycle was in effect.

Apparatus

The experimental chamber was 32 cm along the sides and 45 cm high. The floor was wire mesh and all walls and the ceiling were Plexiglas. The front wall was covered with black paper. The chamber was located in a sound-attenuating box, and white noise was present continuously. A 100-W white houselight, permanently lit, was located on the ceiling of the outer box, 90 cm above the front wall. A one-way mirror parallel to the front panel permitted observation.

The front wall was equipped with two 2.2-cm diameter response keys symmetrically located 2.5 cm to either side of the midline 23 cm above the floor. A force of 0.2 N on either response key operated a microswitch producing an audible click. Each key could be illuminated with a 5-W orange light. Directly below the keys a hopper opening (4.5 by 7 cm and 7 cm from the floor) permitted access to

mixed grain. A 7.5-W white light illuminated the hopper when grain was delivered. All events were controlled and data were recorded by a Commodore® 64 computer.

Procedure

Pretraining. Sessions were conducted daily at approximately the same hour (9:00 a.m.). All pigeons were trained to peck both keys under a modified autoshaping procedure developed by Schwartz (1980). After variable intertrial intervals ($M = 60$ s) one or both keys were lit randomly for 6 s after which reinforcement (4-s access to grain) was delivered. During reinforcement the keylights were turned off. If a peck occurred on an illuminated key, food was presented immediately. Each session ended when 50 reinforcers had been delivered. Autoshaping lasted for four or five sessions.

Before the experiment proper, a transition phase providing intermittent reinforcement for pecks on each key was implemented for four sessions. Each trial began with one key lit randomly. A peck on the lit key turned off the keylights for a 1-s interpeck interval after which one of the keys was again lit randomly. Pecks during the period of darkness reset the interval but produced no other experimentally programmed consequences. The probability of reinforcement after each set of four pecks was gradually decreased from 1 to .4, and the number of trials increased from 50 to 100. All unreinforced trials were followed by a 4-s timeout, during which the keylights were darkened but the houselight remained illuminated. Pecks during the timeout reset the interval but had no other scheduled consequences. Thus, the interpeck interval and timeout differed only in duration. Pecks during the timeout were rare and did not produce systematic differences in the rates of reinforcement.

During the experiment proper, both keylights were illuminated simultaneously, and each peck to either key darkened both for the 1-s interpeck interval. Each sequence of four pecks was given a variability score equal to the number of sequences intervening between the current one and its most recent match (to a maximum of 50). If this variability score was greater than a criterion number, reinforcement was delivered with probability $u = P(\text{S}^+|\text{Cr})$.

All nonreinforced sequences ended in timeout. To obtain the desired probability of a criterional response, $P(\text{Cr})$, the last 20 variability scores were arranged in ascending order and a probability 1/20 was associated with each score. From the *cumulative distribution function* the kth percentile ($k = 1, 2, \ldots, 99$) of variable X (variability scores) was defined as the *largest* x_k that satisfied the inequality

$$P(X < x_k) < (k/100). \qquad (3)$$

In other words, when the 20 sample scores were arranged in ascending order, x_k, the criterion for the next trial, corresponded to the $n = [20(k/100)]$ observation (e.g., the 70th and the 50th percentiles matched the 14th and the 10th observation, respectively). Note that the generally adopted percentile formula (see Galbicka, 1988; Platt, 1973), which would set the kth percentile equal to the $(20 + 1)(k/100)$ observation, was not used because it presumes an essentially continuous response distribution in clear contrast with the discrete nature of variability scores. In addition, computer simulations showed that using (20) instead of (20 + 1) in the percentile formula provided better control over the probability of criterional responding. On the other hand, Equation 3 implies that $P(X > x_k) \leq [1 - (k/100)] = P(\text{Cr})$. On each trial, then, the probability that the next score exceeded the criterion was either equal to, or slightly less than, $P(\text{Cr})$. Whenever the latter occurred, which was due to the presence of ties (more than one observation sharing the criterional score), the following two rules were applied to maintain reinforcement probability constant: (a) Scores strictly exceeding the criterion were reinforced with probability $u = P(\text{S}^+|\text{Cr})$; (b) scores tying the criterion were reinforced with the desired overall probability of reinforcement, $P(\text{S}^+)$. When $P(X > x_k) = P(\text{Cr})$ or, in other words, when exactly $20P(\text{Cr})$ sample observations exceeded the criterional score, only rule (a) was applied. This algorithm always was used except when the criterion was zero. Zero variability scores (i.e., repetitions) were never reinforced, even though they might otherwise have been considered criterional. Previous pilot studies had shown that when repetitions were reinforced intermittently, behavior became highly stereotyped and lost contact with the scheduled contingencies.

Table 1

Experimental conditions for each group and obtained results. Variability scores were based on the last five sessions. $P(S^+) = .3$ for all groups. See text for further details.

Group	Bird	$P(Cr)$	$P(S^+|Cr)$	Session	Reinforcement		Variability scores		
					M^a	SD^a	M	Minimum	Maximum
1	1	.3	1	30	31.7	2.19	14.11	12.75	15.12
	7	.3	1	30	32.0	2.17	14.06	13.22	14.48
	9	.3	1	31	31.7	2.43	13.80	12.68	14.88
	10	.3	1	24	32.2	2.47	13.64	11.90	15.13
2	2	.5	.6	31	31.3	4.40	14.12	13.44	14.59
	5	.5	.6	32	31.6	4.63	13.24	12.23	14.62
	6	.5	.6	30	32.4	4.40	12.30	11.72	12.90
	11	.5	.6	30	31.6	4.62	13.75	13.35	14.30
3	3	.75	.4	31	31.6	4.56	11.70	10.18	12.68
	4	.75	.4	31	31.6	4.87	12.08	11.08	13.15
	8	.75	.4	30	30.9	5.01	7.22	6.20	8.09
	12	.75	.4	29	31.2	5.00	8.10	5.88	10.62
4	13	1	.3	25	28.5	4.76	10.34	8.38	11.12
	14	1	.3	20	28.6	3.66	1.56	1.02	3.14
	15	1	.3	26	29.3	3.54	10.37	8.43	11.32
	16	1	.3	20	30.6	4.10	0.76	0.51	1.20
	17	1	.3	20	30.1	5.53	3.35	3.07	5.13
	18	1	.3	26	29.9	3.80	4.03	2.86	5.67

[a] First session not included.

Another decision concerned the maximum allowed criterion number. Suppose the criterion was set to 30. This meant that the next sequence should differ from at least the preceding 30 patterns in order to be reinforceable. Nonetheless, an impossible situation would occur if all the 16 possible sequences were among the last 30. This impossible situation is most likely if behavior is highly variable and a high percentile is currently being used. To prevent its occurrence, the maximum criterion number was set at 25, because the probability of emitting the 16 sequences on the last 25 trials, assuming random responding, is less than .01 (in fact, it occurred only once throughout the present experiments; in this case the criterion was set to the highest possible score). The slight underestimation of the probability of criterional responses stemming from this limitation was thought to be meaningless.

During the first session, responding on the first trial was reinforced. The criterion during the first 20 trials was based on all previous scores. Even adjusting the criterion as the sample size increased, the probability of a criterional response and, consequently, the probability of reinforcement were underestimated because variability scores, up to the 50th trial,

were also a function of trial number (e.g., only after Trial 10 could variability scores be greater than 10). From the second session onwards, the last 20 variability scores each subject produced in the previous session were loaded into the computer's memory and were used to compute the initial criteria.

For this experiment, 18 pigeons were assigned randomly to four groups (see Table 1). For Groups 1, 2, and 3, the probability of a criterional response was .3, .5, and .75, and for Group 4 it was 1 (i.e., all responses were eligible for reinforcement). The last group permitted an assessment of the effects of reinforcement intermittency in the absence of any variability requirement. In order to have $P(S^+)$ = .3 under all groups, the conditional probability of reinforcement was adjusted to equal 1, .6, .4, and .3 for Groups 1 to 4, respectively. Subjects received from 20 to 32 sessions until mean variability scores showed no clear trends.

RESULTS

Summary statistics about the number of reinforced trials obtained by each bird during all sessions, excluding the first, are shown in Table 1. Means were close to, but slightly above, 30, the expected value for all groups.
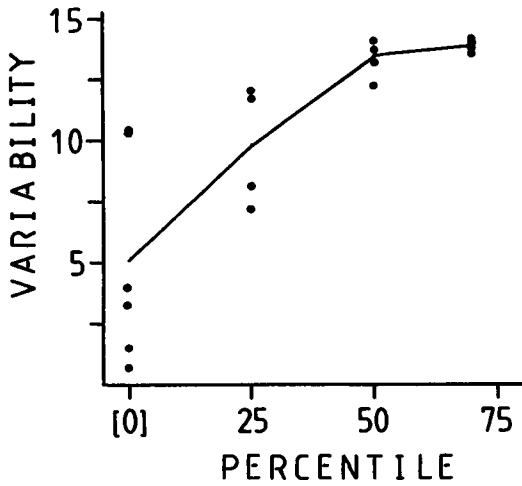
Fig. 2. Mean variability score as a function of the percentile. (Dots = averages over the last five sessions for each subject; solid line connects the mean of each group.) Variability was permitted but not required in Group 4, corresponding to Percentile 0.
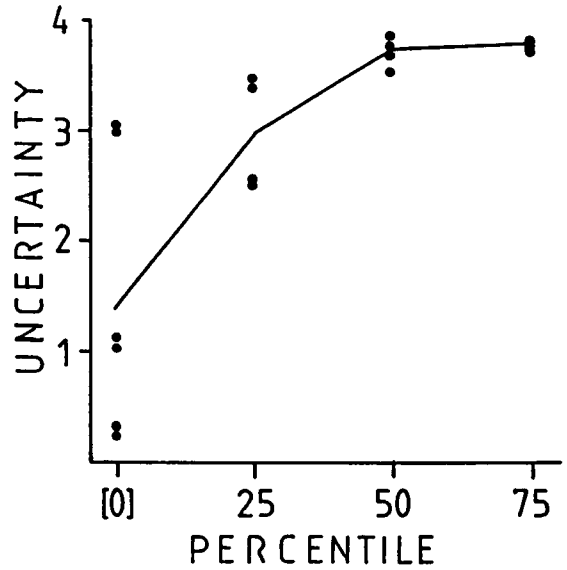

Fig. 3. Uncertainty as a function of the percentile. (Dots represent averages over the last five sessions for each subject; the solid line connects the mean of each group.)

Standard deviations were also similar in Groups 2, 3, and 4, but smaller in Group 1. The magnitude of these anomalies was not large enough to undermine the analysis that follows.

The mean variability score each bird obtained in the last five sessions (taken as one long session of 500 trials) is shown in Figure 2. Note that data are arranged by groups with each percentile corresponding to a different group of birds. Individual values are shown in Table 1 together with the minimum and the maximum mean variability score obtained during any one session. Random responding would produce a mean variability score of 15, the expected number of trials before repeating a particular pattern. An ANOVA by ranks (see Meddis, 1984) supported the idea that the amount of variability depended on the requirement—$Z = 3.63$, $p < .001$ (specific alternative hypothesis, H1: M4 < M3 < M2 < M1, where M4 means expected mean rank of Group 4, etc.).

Another useful measure of sequence variability is uncertainty (U), derived from information theory (see Attneave, 1959; Miller & Frick, 1949; Page & Neuringer, 1985). Uncertainty was computed according to the following equation:

$$U = - \sum_{i=1}^{16} [p_i \log_2(p_i)] , \qquad (4)$$

where U is uncertainty measured in bits and p_i stands for the probability of sequence i. Uncertainty is maximal, and equal to four bits, when the 16 patterns are all equally likely. Uncertainty is minimal, and equal to zero bits, if only one sequence is emitted. Figure 3 shows mean U values in the last five sessions. The trend was the same as in Figure 2; that is, group means were a negatively accelerated function of the requirement. Taken together, these findings are consistent with the results reported by Page and Neuringer (1985).

Within-group differences were another important aspect of the data. The magnitude of these differences was inversely related to the requirement (see Figures 2 and 3). Even when no variability was demanded, Birds 13 and 15 of Group 4 still maintained a certain amount of variation, whereas the modal sequences for other subjects in this group represented exclusive responding on one key, accounting for from 80% to 92% of their behavior. Page and Neuringer (1985, Experiment 5) also observed greater intersubject differences when the variability requirement was totally eliminated.

Birds 3 and 4 of Group 3 also maintained a high degree of variability in the presence of a very low requirement. Birds in this group showed a different kind of stereotypic behavior. For Bird 12 sequences LLLL and RLLL accounted for 60% of its behavior; for Bird 8, modal sequences involved no more than one switch between keys (patterns LLLL, RLLL, RRLL, and RRRL accounted for 86% of the sequences; this bird scarcely switched from the left to the right key); for Bird 3, stereotypy was exhibited on the last peck of the sequence (79% of these pecks on the right key and only 21% of the left key). On the other hand, Groups 1 and 2 showed a very small spread in individual subjects' data. The only stereotypical feature in Group 2 was a preference, in Birds 5, 6 and 11, for the right key at the beginning of each sequence (from 61% to 63%). Birds in Group 1 generated highly unpredictable behavior. As an example, Figure 4 presents the frequency of each variability score for Bird 1 and the predicted values assuming random responding (see Equation 1). A χ^2 test showed no significant difference between the two distributions, $\chi^2(36) = 37.44, p > .40$.

DISCUSSION

The percentile reinforcement schedule performed appropriately. The probability of a criterional response was, in each case, close to the predicted value and, consequently, the probability of reinforcement was handled adequately. Several factors might account for the remaining minor deviations: the previously mentioned underestimation of response probability for Group 1, the sample size from which the criterion was calculated, the algorithm deciding what do to when ranks tied (see Procedure), and even the computer's pseudorandom number generator. The slight underestimation of response probability for Group 1 may be corrected by letting the maximum criterion number be determined on a trial-by-trial basis. More specifically, a check could be made continuously to see if the actual criterion, corresponding to the 70th percentile, was really attainable, that is, if it was possible to emit a sequence whose variability score would equal or exceed the criterion number. In this case, the criterion could sometimes be greater than 25 and still be a valid criterion.
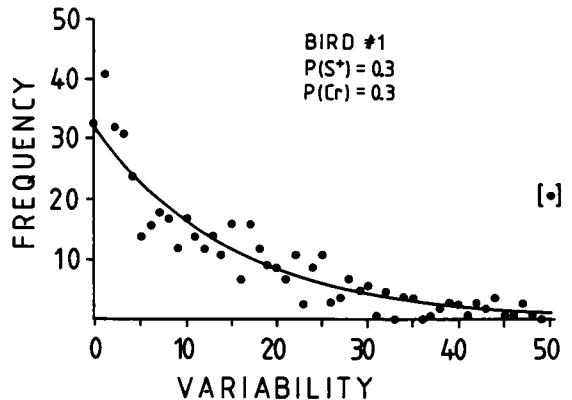
Using a between-subjects design, the pres-

Fig. 4. Dots represent the frequency distribution of variability scores obtained with Bird 1 under the last five sessions. The last dot corresponds to scores greater than 49. The solid line connects the expected frequencies assuming random responding (see Equation 1).

ent experiment demonstrated that different amounts of behavioral variability can be maintained by operant contingencies. With the exception of Page and Neuringer's (1985) study, no other research had addressed this important subject. In fact, an adequate theory of learning must decide whether to take behavioral variability as a fundamental behavioral dimension, directly amenable to reinforcement, or as a secondary behavioral property, reducible to more basic processes (see Neuringer, 1986; Page & Neuringer, 1985; Schwartz, 1982). This is both an empirical and a theoretical question. If reinforcement is seen as strengthening behavior (i.e., as increasing the probability of those responses that produce it), how can this process engender variability? This theoretical question is dealt with in the General Discussion. In contrast, if variability is a by-product of more fundamental processes, what are these processes? The present experiment, together with Page and Neuringer's Experiment 5, ruled out differences in reinforcement intermittency as the explanation of the differences in the asymptotic level of variability, but only further research can settle the issue.

The high degree of pattern variability maintained by some birds in the presence of a weak requirement (Group 3) parallels Galbicka and Platt's (1989) finding of weak contingencies generating good differentiation if these contingencies are obtained by reinforcing criterional

responses intermittently and never reinforcing noncriterional responses. On the other hand, the high degree of variability maintained by Birds 13 and 15 in the absence of any requirement whatsoever resembles the results found with other response dimensions when reinforcement ceases to be response dependent (e.g., Davis & Platt, 1982). In studies in which dimensions other than variability (e.g., response location) are differentiated, increased variability generally means decreased control by the previously differentiated dimension; in the present study increased variability means precisely the opposite, that is, it reflects maintenance of control. In other words, pattern stereotypy in the present investigation is analogous to response variability in these other studies, and vice versa.

A plausible hypothesis to account for these data is to argue that behavior is variable initially; whether it remains variable or not is a matter of how stringent the current variability requirements are. With weaker requirements different variables might control the performance of different birds; with stronger requirements the influence of these other variables is greatly attenuated, and the requirement per se plays the major role. Consequently, weaker requirements predict greater intersubject differences (see Crow, 1977, and Neuringer, 1986, for related hypotheses). Scott and Platt (1985, Experiment 4) using a percentile schedule in which reinforcer delivery was contingent on the location of a joystick displacement response emitted by rats, also found greater within-group variance when the requirement to deliver reinforcement was less stringent. Further research should clarify this topic; a possible starting point is the replication of the present experiment with a within-subjects design and an initially stereotyped baseline.

EXPERIMENT 2

Changes in reinforcement probability have well-known if somewhat inconsistent effects on behavioral variability (e.g., Antonitis, 1951; Boren, Moersbaecher, & Whyte, 1978; Eckerman & Lanson, 1969; Ferraro & Branch, 1968; Herrick & Bromberger, 1965; Herrnstein, 1961; Lachter & Corey, 1982; Millenson, Hurwitz, & Nixon, 1961; Notterman & Mintz, 1965; Stebbins & Lanson, 1962; Tre-

mont, 1984; for a review, Boulanger et al., 1987). However, no research has been conducted to study the effects of changes in reinforcement probability when variability is the response property on which reinforcement is dependent. In fact, variability due to reinforcement intermittency might interact with operant variability, a function of an operant contingency. The percentile reinforcement schedule, as previously defined, is well suited to study this interaction because its main feature is the relative dissociation it provides between the criterion defining the operant class (the percentile) and the overall probability of reinforcement.

When only criterional responses are eligible for reinforcement, the maximum probability of reinforcement is, obviously, the probability of these criterional responses (i.e., the complement of the percentile). This means that substantial manipulation of reinforcement probability can occur only if a low percentile is chosen. In the present experiment the 30th percentile defined criterional responses and the overall probability of reinforcement was systematically varied from .3 to .7.

METHOD

Subjects and Apparatus

Eight experimentally naive pigeons were assigned randomly to two groups of 4 birds each. The apparatus was the same as in Experiment 1.

Procedure

Autoshaping was as in Experiment 1. The 30th percentile was used throughout, so $P(Cr) = .7$. For Group LH (L = low, H = high), $P(S^+|Cr)$ was initially set to .43 and then to 1. Hence, $P(S^+)$ changed from $.3 = (.7)(.43)$ to $.7 = (.7)(1)$. For Group HL the reverse order occurred (i.e., $P(S^+) = .7$ and then $P(S^+) = .3$). Table 2 shows the number of sessions for each condition. The transition phase (see above, Experiment 1, Procedure) was implemented only for Group LH. All other details were the same as for Experiment 1.

RESULTS

Table 2 shows the mean and standard deviation of the number of reinforced trials obtained by each bird on both conditions. These values closely matched predicted ones but some deviations are worth mentioning. The number

Table 2

Experimental conditions and results of Experiment 2. $P(Cr) = .7$ throughout.

| Group | Bird | $P(S^+) = .3$[a] | | | | | $P(S^+) = .7$[a] | | | |
| | | | Reinforcement | | Mean variability score | | | Reinforcement | | Mean variability score |
		Session	M[b]	SD[b]		Session		M	SD	
LH	1	26	29.0	2.86	13.38	29		68.4	4.04	7.43
	3	30	25.4	5.15	10.20	25		69.4	2.71	12.92
	4	27	28.5	3.37	13.76	15		69.9	2.49	13.58
	8	27	26.3	5.64	13.31	20		69.9	2.10	13.28
HL	2	25	31.7	5.51	12.01	27		68.8	3.62	11.87
	5	20	30.1	4.97	9.46	28		69.4	2.88	9.91
	6	20	30.4	3.49	9.70	29		67.2	6.26	13.15
	7	25	30.8	5.13	12.60	24		68.8	3.28	13.84

[a] For Group LH, $P(S^+)$ was initially .3 and next .7. For Group HL, the reverse order occurred, that is, $P(S^+) = .7$ and later $P(S^+) = .3$.

[b] First session not included in Group LH when $P(S^+) = .3$.

of reinforced trials was too low in some sessions of Birds 3 and 8 when $P(S^+) = .3$ and of Bird 6 when $P(S^+) = .7$.

Mean variability scores in the last five sessions (see Table 2) showed no consistent trend when reinforcement probability was changed markedly. Some birds did not significantly alter their behavior (Birds 4 and 8 from Group LH; Birds 2 and 5 from Group HL). For Birds 6 and 7 from Group HL, decreasing reinforcement probability reduced response variability. In Group LH, increasing reinforcement probability increased response variability in Bird 3 and decreased it in Bird 1.

The mean variability scores obtained when $P(S^+) = .3$ and $P(S^+) = .7$ were subjected to an ANOVA by ranks for related samples, with subject as a blocking variable. No significant difference was found ($K = 12, p > .05$). Mean uncertainty scores yielded similar results.

Stereotypical patterns were similar to those observed in Experiment 1. As an example, Figure 5 shows, for Bird 1, U values for each key peck within a sequence, namely, the uncertainty of the first, second, third, and fourth pecks, considered as distinct events. In each case, only two possibilities exist, a left or a right peck. Hence, according to Equation 2, U is maximal, and equal to 1, when $P(L) = P(R) = .5$, and $U = 0$ when $P(L) = 1$ or $P(L) = 0$.

It can be seen that, in Sessions 4, 5, 10, 18 and 20, when $P(S^+) = .3$, a systematic preference for one response key was clearly shown on the first peck ($U1 < .8$), all successive pecks

remaining generally highly variable. A different picture was visible when $P(S^+)$ changed to .7 (after Session 25 in Figure 5). Variability on the first peck gradually decreased until, in Session 37, 99 of the 100 sequences started on the left key ($U1 = .08$). In Session 39, the second peck was also highly stereotyped ($U2 = .70$). From then on, a different pattern emerged whereby variability was confined to the first peck, all successive pecks being emitted frequently on the same key (on the last session $U1 = .87$, $U2 = .56$, $U3 = .44$, and $U4 = .37$). In other words, when $P(S^+) = .3$, variability increased from the first to the last peck; in the last sessions with $P(S^+) = .7$, the reverse occurred, that is, variability decreased from the first to the last peck.

Other birds (e.g., Bird 2 when $P(S^+) = .3$ and Bird 5 when $P(S^+) = .3$ as well as when $P(S^+) = .7$) maintained a strong preference for one particular key on the first and/or second peck of each sequence. Switching stereotypy was also displayed by Bird 3 when $P(S^+) = .3$ (sequences LLLL, LLLR, LLRR (23%), LRRR (23%), and RRRR accounting for 72% of its behavior). No conclusion could be drawn relating probability of reinforcement to kind of stereotypy.

DISCUSSION

Substantial changes in reinforcement probability did not produce consistent changes in the level of variability. This finding is not completely unexpected because it parallels the results obtained when variability is not the re-
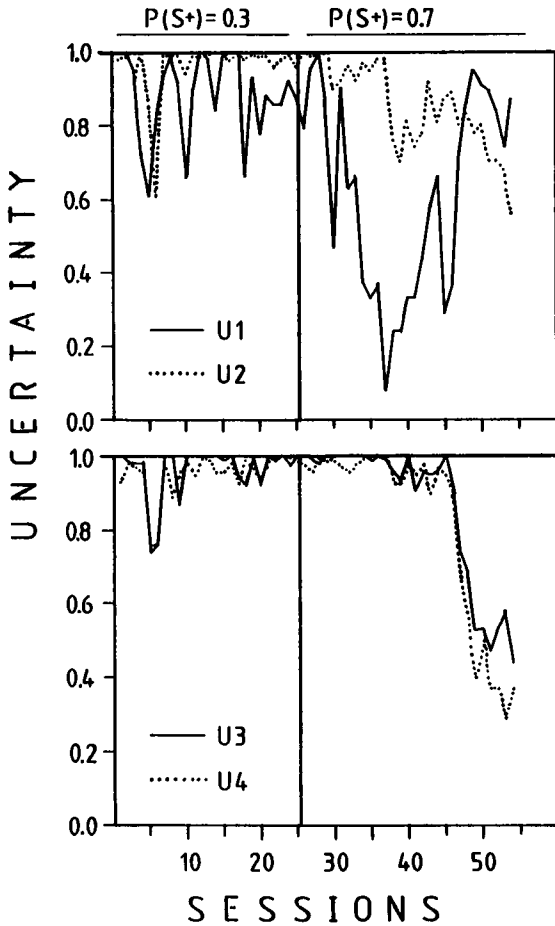
Fig. 5. Uncertainty of each key peck obtained with Bird 1 when $P(S^+)$ was .3 and .7. $U1, \ldots, U4$ stands for uncertainty of the first, second, third, and fourth pecks within a sequence. See text for further details.

sponse dimension defining the operant class. Whereas several authors found an increase in variability when reinforcement became less frequent (Eckerman & Lanson, 1969; Ferraro & Branch, 1968; Herrick & Bromberger, 1965; Lachter & Corey, 1982; Stebbins & Lanson, 1962; Tremont, 1984), others did not find a similar result (Boren et al., 1978; Herrnstein, 1961; Millenson, et al., 1961; see Boulanger et al., 1987, for possible interpretations of these inconsistent findings). In the present situation, several factors might have obscured the effects of reinforcement probability. One of them is the timeout that may have gained discriminative control over the initial peck of some sequences. Due to its contiguity with timeout, the location of the first peck of a reinforced

sequence could more easily be remembered and hence repeated; consequently, more sequences beginning with the first peck on the same key would be reinforced, originating a positive feedback loop. This process might account for the increased stereotypy of Bird 1 when $P(S^+)$ = .7 (see Figure 5, $U1$). That timeout cannot be the only variable at play is suggested by the presence of other types of stereotypy (e.g., stereotypy of the *last* three pecks).

The percentile schedule presented some limitations when performance became highly stereotyped; when the probability of a repetition exceeded the probability associated with the chosen percentile, .3, the criterion variability score remained at zero, but the probability of a criterional response was less than .7 = 1 − .3 because only variability scores strictly exceeding zero were eligible for reinforcement. On these occasions, due to the decision to treat repetitions independently, a positive correlation existed between variability and reinforcement probability. This correlation was maintained until the probability of a repetition was less than the percentile value. One solution to this problem is to increase the number of pecks making up a sequence such as in Page and Neuringer's (1985) study in which eight pecks per sequence were used. When this happens, and assuming the probability of a right (or left, respectively) key peck is not changed, the probability of a repetition is reduced considerably (e.g., assuming $P(R) = P(L) = .5$ and no sequential dependencies in responding, the probability of a repetition is halved each time the numbers of pecks per sequence are increased.).

GENERAL DISCUSSION

The present experiments developed and tested a new percentile reinforcement schedule suited to study pattern variability. The main feature of the schedule is the dissociation between overall probability of reinforcement and the criterion defining the operant class. This enables the experimenter to manipulate reinforcement probability and the variability requirement independently. Despite some deviations from predicted performance, the schedule successfully achieved its main purposes. In Experiment 1, the variability requirement was manipulated between groups while the overall probability of reinforcement

was held approximately constant. Results showed that a weak contingency only moderately controls the amount of generated variability. Thus, a large change in reinforcement probability could, theoretically, have shown its effects during Experiment 2. However, no consistent effect could be ascribed to reinforcement probability per se or, in other words, to the effects of intermittent reinforcement when an operant contingency was simultaneously present. Other variables, such as the timeout, might have competed with probability of reinforcement, obscuring any effects this variable might have. Despite important within-group differences, results from Experiment 1 showed that variability directly increased with the requirement, a conclusion already reported by Page and Neuringer (1985). This finding raises a basic question previously addressed by several authors (e.g., Page & Neuringer, 1985; Schwartz, 1982) and briefly alluded to earlier: How can reinforcement maintain behavioral variation if it increases the probability of the class of behaviors producing it? Another way to state the question is "What objective property of responses [upon which reinforcement is dependent] would unite them into a class?" (Schwartz, 1982, p. 178). It could be argued that, on each trial, a certain number of sequences, if generated, would yield variability scores exceeding the criterion. Hence, "being different from the *N* previous patterns" was the defining property on which reinforcement was contingent. However, the question asked by Schwartz is about behavioral processes, not about procedures (see Catania, 1973) and the answer assumes that pigeon's behavior is somehow sensitive to the dimension "being different from *N* previously emitted patterns." This assumption is most unlikely given the difficulty pigeons have remembering sequences beyond the trivial LLLL and RRRR patterns.

Page and Neuringer (1985) and Neuringer (1986) circumvented the question raised by Schwartz by considering variability a fundamental behavioral dimension, such as force or duration, and by stressing the *shaping* in place of the *strengthening* effects of reinforcement: "When the experimenter shapes keypecking . . . the pigeon is taught where, when, and possibly how fast or hard, and so on, to peck. Analogously, there may be a dimension of all behaviors, described as *variability,* with which the organism enter our experiments . . . Turn-

ing on or off a variability generator may be under the control of reinforcement, but the variability generator is not itself created through reinforcement" (Page & Neuringer, 1985, p. 450). The relationship between level of variability and degree of requirement is further viewed as a fine tuning of this inborn variability generator by the current contingencies of reinforcement.

Whatever the adequacy of this conceptualization, the hypothesis of a more basic behavioral process underlying the generation of behavioral variability has not been fully appreciated. Intermittent reinforcement is one possibility, but the present results did not support it. An alternative view might ascribe the operant conditioning of response variability to a process of *probability-dependent selection.* Results of Experiment 1 suggested that overall probability of reinforcement does not play the major role in the process of variability differentiation. However, we are still left with the (conditional) probability of reinforcement per pattern. Consider what happens to the probability of reinforcement—$P(s^+)$—associated with a particular sequence after it has been emitted. Whatever the requirement, $P(s^+) = 0$ on the next trial because no repetition is ever reinforced. When a low percentile is chosen, after a few trials without emitting the sequence, $P(s^+)$ equals $P(S^+|Cr)$, whereas when a high percentile is used, more trials have to elapse before the pattern is considered criterional again, and $P(s^+) = P(S^+|Cr)$. On every trial, then, the more stringent the requirement the smaller the subset of sequences that are eligible for reinforcement. Also, this smaller subset is more likely restricted to the momentarily least probable patterns. Stated differently, increasing percentile values assures *differential reinforcement of increasingly less probable patterns.* A similar idea is found in Blough's (1966) least frequent IRT reinforcement schedule.

If this interpretation is valid, variability should not be considered a fundamental (i.e., an irreducible) behavioral dimension, but rather as the outcome of more primary processes, such as probability-dependent selection. This hypothesis predicts random responding as the asymptotic performance when stronger and stronger requirements are used. In the limiting case, all patterns are equally likely and, consequently, equally reinforced.

REFERENCES

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42,* 273-281.

Attneave, F. (1959). *Applications of information theory to psychology.* New York: H. Holt.

Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior, 9,* 581-591.

Boren, J. J., Moersbaecher, J. M., & Whyte, A. A. (1978). Variability of response location on fixed-ratio and fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 30,* 63-67.

Boulanger, B., Ingebos, A., Lahak, M., Machado, A., & Richelle, M. (1987). Variabilité comportementale et conditionnement opérant chez l'animal. *L'Anneé Psychologique, 87,* 417-434.

Bryant, D., & Church, R. M. (1974). The determinants of random choice. *Animal Learning & Behavior 2,* 245-248.

Catania, A. C. (1973). The concept of the operant in the analysis of behavior. *Behaviorism, 1(2),* 103-116.

Crow, L. T. (1977). Is variability a unifying behavioral concept? *Psychological Record, 27,* 783-790.

Davis, E. R., & Platt, J. R. (1983). Contiguity and contingency in the acquisition and maintenance of an operant. *Learning and Motivation, 14,* 487-512.

Eckerman, D. A., & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior, 12,* 73-80.

Ferraro, D. P., & Branch, K. H. (1968). Variability of response location during regular and partial reinforcement. *Psychological Reports, 23,* 1023-1031.

Galbicka, G. (1988). Differentiating *The Behavior of Organisms. Journal of the Experimental Analysis of Behavior, 50,* 343-354.

Galbicka, G. (in press). Shaping and response differentiation: Creating response units through differential selection. In W. F. Angermeier & S. R. Hursh (Eds.), *Grundlagen des operanten verhaltens.* Munich: Reinhardt-Verlag.

Galbicka, G., & Platt, J. R. (1989). Response-reinforcer contingency and spatially defined operants: Testing an invariance property of phi. *Journal of the Experimental Analysis of Behavior, 51,* 145-162.

Herrick, R. M., & Bromberger, R. A. (1965). Lever displacement under a variable ratio schedule and subsequent extinction. *Journal of Comparative and Physiological Psychology, 59,* 392-398.

Herrnstein, R. J. (1961). Stereotypy and intermittent reinforcement. *Science, 133,* 2067-2069.

Lachter, G. D., & Corey, J. R. (1982). Variability of the duration of an operant. *Behaviour Analysis Letters, 2,* 97-102.

Meddis, R. (1984). *Statistics using ranks: A unified approach.* Oxford: Blackwell.

Millenson, J. R., Hurwitz, H. M. B., & Nixon, W. L. B. (1961). Influence of reinforcement schedules on response duration. *Journal of the Experimental Analysis of Behavior, 4,* 243-250.

Miller, G. A., & Frick, F. C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review, 56,* 311-324.

Morris, C. J. (1987). The operant conditioning of response variability: Free-operant versus discrete-response procedures. *Journal of the Experimental Analysis of Behavior, 47,* 273-277.

Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. *Journal of Experimental Psychology: General, 115,* 62-75.

Notterman, J. M., & Mintz, D. E. (1965). *Dynamics of response.* New York: Wiley.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 429-452.

Platt, J. R. (1973). Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in theory and research* (Vol. 7, pp. 271-296). New York: Academic Press.

Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12,* 653-661.

Schoenfeld, W. N., Harris, A. H., & Farmer, J. (1966). Conditioning response variability. *Psychological Reports, 19,* 551-557.

Schwartz, B. (1980). Development of complex, stereotyped behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 33,* 153-166.

Schwartz, B. (1982). Failure to produce response variability with reinforcement. *Journal of the Experimental Analysis of Behavior, 37,* 171-181.

Scott, G. K., & Platt, J. R. (1985). Model of response-reinforcer contingency. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 152-171.

Stebbins, W. C., & Lanson, R. N. (1962). Response latency as a function of reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 5,* 299-304.

Tremont, P. J. (1984). Variability of force and interresponse time under random interval reinforcement schedules. *Behavioural Processes, 9,* 413-420.