Variability Is an Operant

Article *in* Journal of Experimental Psychology Animal Behavior Processes · July 1985

DOI: 10.1037/0097-7403.11.3.429

2 authors:

Suzanne Page
LIRIC
8 PUBLICATIONS **619** CITATIONS
SEE PROFILE

Allen Neuringer
Reed College
87 PUBLICATIONS **4,743** CITATIONS
SEE PROFILE

Journal of Experimental Psychology:
Animal Behavior Processes
1985, Vol. 11, No. 3, 429–452

0097-7403/85/$00.75

Variability Is an Operant

Suzanne Page and Allen Neuringer
Reed College

Pigeons were rewarded if their pattern of eight pecks to left and right response keys during the current trial differed from the patterns in each of the last *n* trials. Experiments 1 and 2 compared Schwartz's (1980, 1982a) negative findings (variability was not controlled by reinforcement) with the present positive results and explained the difference. Experiment 3 manipulated *n* and found that the pigeons generated highly variable patterns even when the current response sequence had to differ from each of the last 50 sequences. Experiment 4 manipulated the number of responses per trial; variability increased with increasing responses per trial, indicating that the pigeons were acting as quasi-random generators. Experiment 5 showed that for high levels of variability to be engendered, reinforcement had to be contingent on response variability. In a yoked condition, where variability was permitted but not required, little response variability was observed. Experiment 6 demonstrated stimulus control: Under red lights the pigeons generated variable patterns, and under blue lights they repeated a particular fixed pattern. We concluded that behavioral variability is an operant dimension of behavior controlled by contingent reinforcement.

Is response variability controlled by contingent reinforcers, as are other behavioral dimensions, such as response rate, location, duration, force, and topography? That is, can behavioral variability be increased or decreased by reinforcers contingent on such increases or decreases? Variability is necessary for many behavioral phenomena. The process of operant shaping depends on a variable substrate (Skinner, 1938). Successive approximations to some goal response are selected for reinforcement, and without sufficient variation, selection is difficult or impossible. Behavioral variability is also important for problem solving and creativity. It would be useful to know, in these instances, whether variability is controlled by its consequences. The question also has theoretical importance. Reinforcers are said to increase the probability of those specific responses that produce them. Thus, reinforcement might inexorably lead to response repetition and therefore to decreased variability. If variability can be reinforced, what function does reinforcement serve?

Schedules of reinforcement clearly affect behavioral variability (Antonitis, 1951; Crossman & Nichols, 1981; Eckerman & Lanson, 1969; Herrnstein, 1961; Lachter & Corey, 1982; Notterman, 1959; Piscaretta, 1982; Schwartz, 1980, 1981, 1982a, 1982b). Intermittent schedules, which reinforce only occasional responses, generally engender higher variability along many dimensions than does reinforcement of every response (e.g., Lachter & Corey, 1982). However, these demonstrations are of respondent effects and are therefore orthogonal to the present question. An analogous case would be to reward rats for running in a running wheel. One might observe changes in the rat's heart rate, but reinforcement was not contingent on heart rate. Whether heart rate could be manipulated through contingent reinforcement is a separate issue (Miller, 1978). So, too, the question of whether behavioral variability can be reinforced is independent of schedule-eliciting effects.

There have been few attempts to reinforce variability directly, and the results are inconsistent. Pryor, Haag, and O'Reilly (1969)

This article was derived from an undergraduate thesis submitted by the first author to Reed College. We thank Scott Gillespie, Rick Wood, Steve Luck, and Richard Crandall for invaluable technical advice and assistance and Barry Schwartz for helpful discussion and suggestions. Order of authorship was determined by a quasi-random process.

Requests for reprints should be sent to Allen Neuringer, Department of Psychology, Reed College, Portland, Oregon 97202.

reinforced novel behaviors in two porpoises. Reinforcers were delivered for any movement that in the opinion of the trainer did not constitute normal swimming motions and that had not been previously reinforced. Different types of behaviors emerged, including several that had never before been observed in animals of that species. In a more controlled setting, Schoenfeld, Harris, and Farmer (1966) delivered rewards to rats only if successive interresponse times fell into different temporal class intervals. This procedure produced a low level of variability in the rat's rate of bar pressing. In the most sophisticated and demanding experiment to date, Blough (1966) reinforced the least frequent of a set of interresponse times and obtained interresponse time distributions in three pigeons that approximated a random exponential distribution. Similarly, Bryant and Church (1974) rewarded rats 75% of the time for alternating between two levers and 25% of the time for repeating a response on the same lever and found that for 3 of 4 subjects, the resulting choice behaviors could not be distinguished from chance. People have also been rewarded for behaving variably in more natural environments (Holman, Goetz, & Baer, 1977).

Although the above studies indicate that variability can be reinforced, the conclusion is not secure. When the response in question is a complex behavior, such as the responses studied by Pryor et al. (1969) and Holman et al. (1977), the observer plays an important role in defining *novel response,* and the novelty observed might be as much a function of the observer as of the subject. Moreover, Schwartz (1982b) has argued that the Pryor et al. findings with porpoises can be attributed to the effects of repeated extinction and reconditioning, with extinction occurring increasingly rapidly in progressive instances, and not to the direct reinforcement of novelty. Because extinction increases variability (e.g., Antonitis, 1951), the variability observed may have been a by-product of the reinforcement schedule. A similar argument can be made with respect to both Blough (1966) and Bryant and Church (1974). Although unlikely, it is possible that the reinforcement schedules somehow elicited interresponse time variability in the Blough experiment and stay–alternate variability in the work of Bryant and Church. Here, too, the argument is that variability was not reinforced (variability is not an operant) but the schedules of reinforcement engendered variability as a respondent by-product. The control experiments necessary to distinguish between the operant and respondent alternatives were not performed.

The most serious evidence against variability as an operant dimension of behavior comes from Schwartz (1980, 1982a). Pigeons were required to generate on two response keys a pattern of eight responses that differed from the last sequence of eight responses. A 5 × 5 matrix of lights provided the pigeons with continual feedback concerning their distribution of pecks. The light started in the upper left-hand corner; each left-key peck (L) moved the light one column to the right and each right-key peck (R) moved the light down one row. In preliminary training, pigeons were simply required to move the light from upper left to lower right. There was no requirement to vary the pattern. If the birds responded more than four times on either key (thereby moving the light off the matrix), the trial terminated without reinforcement. The pigeons were highly successful, obtaining 70% to 90% of available reinforcers by generating repetitive and energy-conserving patterns (e.g., LLLLRRRR). In the experiments of most concern to the present work (Schwartz, 1980, Experiment 4; 1982a, Experiment 1), in addition to requiring four pecks on each key, the schedule demanded that each eight-response sequence differ from the immediately preceding sequence. Now, when the pigeons had to vary their response patterns to be rewarded, only 40% of the available reinforcers were earned. Number of different sequences emitted, an index of learned variability, did not appreciably increase.

The Schwartz findings, though apparently robust, appeared to conflict with those of Pryor et al. (1969), Blough (1966), and Bryant and Church (1974). Furthermore, Neuringer (1984, 1985) showed that when provided with feedback, people learned to generate highly variable response sequences, indeed sequences so variable that they could be described as random. To explore the apparent disagreement between Schwartz's work and these findings, the present research attempted to reinforce variability in an environment similar to that used by Schwartz, explain Schwartz's negative results, and demonstrate that the variability dimension of behavior is reinforceable and, as with other reinforceable dimensions, sensitive to discriminative stimulus control.

Experiment 1: Variability Versus Variability-Plus-Constraint

The rationale for Experiment 1 derives from the query, How frequently would a random generator be rewarded under the Schwartz (1980, Experiment 4; 1982a, Experiment 1) contingencies where a sequence of eight responses had to differ from the previous sequence for reinforcement? If the Schwartz contingencies reinforced response variability, as was intended, many reinforcers would be given to a random response generator. However, we found that a computer-based random number generator, programmed to respond randomly *left* or *right* under Schwartz's contingencies, was rewarded on only 29% of the trials. This random generator was successful on somewhat fewer trials than were Schwartz's pigeons. The random generator's performance can be explained on the following theoretical grounds. The number of distinct eight-response sequences, given two alternatives, is 256, or 2^8. The number of such sequences meeting the requirement of no more than four responses on each key is 70. Therefore, the probability that a sequence meets the no-more-than-four requirement is 0.27, or approximately the percentage of times that the random generator was successful. According to this reasoning, the no-more-than-four requirement was responsible for the random simulator's low success rate. To determine whether these theoretical considerations can explain Schwartz's findings, the present experiment compared Schwartz's schedule with one in which the requirement of no more than four responses per key was omitted.

Method

Subjects

Three experimentally naive pigeons (Nos. 28, 68, and 71) and 1 with previous experience in conditioning experiments (No. 58) were housed in individual 28-cm × 32-cm × 31-cm cages with grit and water freely available, and they were maintained at approximately 80% of their free-feeding body weights. A 12-hour light-dark cycle was in effect. Sessions were conducted daily unless a bird's weight was more than 20 g over 80% of free-feeding weight.

Apparatus

The experimental chamber measured 28 cm along the sides, 30 cm along the front and rear walls, and 29.5 cm high. The front and rear walls were aluminum, one side wall was pressed board, the other was Plexiglas, and the floor and ceiling were wire mesh. The chamber was housed in a sound-attenuating box equipped with a fan for ventilation and masking noise. A one-way mirror in the outer box, parallel to the Plexiglas chamber wall, permitted observation.

The front wall contained two Gerbrands pigeon keys, each 2 cm in diameter, that were operated by withdrawal of an applied force. The keys were 21.5 cm above the floor; one key was directly above a 4.5-cm × 5.5-cm hopper opening that was itself 7.5 cm above the floor and centered on the panel, and the second key was 4.5 cm (center to center) to the right of the first. A Gerbrands food magazine delivered mixed pigeon grain through the hopper opening. Each key could be transilluminated with a 7.5-W white light, and a 7.5-W white house light was located above the mesh ceiling, directly above the center key. The house light was continuously illuminated except during reinforcement, when the hopper was lighted by a 7.5-W white light.

A 5 × 5 matrix of yellow 5-W cue lights, each 0.75 cm in diameter, was located along the wall to the left of the response keys. The last column of the matrix was 4 cm from the front wall. Columns were 2.5 cm, center to center, and rows were separated by 2 cm, center to center. Each of the 25 lights could be separately illuminated.

Stimulus events were controlled and responses were recorded by a Commodore VIC-20 computer through an integrated circuit and relay interface. Data were recorded on cassette tape at the end of each session and were later transferred to a Digital Equipment PDP-1170 computer for analysis.

Procedure

Pretraining. All pigeons were trained to peck the keys under a modified autoshape procedure derived from Schwartz (1980). After variable intertrial intervals (keys dark) averaging 40 s, one of the two keys, randomly selected, was illuminated for 6 s, followed by presentation of grain reinforcement for 4 s. Pecks to a lighted key immediately produced the reinforcer and darkened the key. Autoshape training continued for four to six sessions after the first session in which a pigeon pecked both lighted keys. Each session terminated after the 60th reinforcer.

A comparison of two main conditions followed. In the variability (V) condition, a sequence of eight left and right responses had to differ from the previous sequence for reinforcement. This was compared with a variability-plus-constraint (VC) condition, where, in addition, exactly four responses had to occur on each key, a condition analogous to Schwartz (1980, Experiment 4; 1982a, Experiment 1).

Variability: Lags 1 and 5. Each trial consisted of eight pecks distributed in any manner over the two keys; each session comprised 50 trials. At the start of each trial, both keys were illuminated with white light. A peck to either key immediately darkened both key lights and initiated a 0.5-s interpeck interval. Pecks during the interpeck interval reset the interval but were not counted and had no other consequence. The eighth peck to a lighted key terminated the trial with either a reinforcer (3.5 s of access to mixed grain) or a 3.5-s time-out, during which the keys were darkened but the house light remained illuminated. Pecks during the time-out reset the 3.5-s timer but had no other consequence. The only difference between time-out and interpeck interval was the duration of these events. Under the Lag 1 condition, whether a trial ended in reinforcement or time-out depended on whether the sequence of eight pecks on that trial differed from the last sequence of eight pecks, that is, whether the response pattern on trial *n* differed from trial $n - 1$. If the sequences on the two trials differed, a reinforcer was given; if the sequences were identical, time-out occurred. Thus, for example, if on Trial 10 the bird pecked LRLLLRRL, Trial 11 would be reinforced unless the bird repeated that sequence. The three experimentally naive birds (Nos. 28, 68, and 71) received 5 to 13 sessions of this Lag 1 training. Because most trials ended with a reinforcer (the pigeons successfully varied their sequences), greater variability was demanded by increasing the look-back value to Lag 5. Now, for reinforcement, the response sequence on trial *n* had to differ from each of the sequences on trials $n - 1$ through $n - 5$, inclusive; otherwise time-out occurred. The one previously experienced bird (No. 58) began the experiment with Lag 5 contingencies. Subjects received 15 to 18 sessions of Lag 5 until the percentage of reinforced trials remained stable over at least 5 sessions. The matrix lights were not used under these variability procedures.

Variability-plus-constraint: Lag 1. Except where noted, the contingencies and parameters here were the same as in the Lag 1 variability condition. As in Schwartz (1980, Experiment 4; 1982a, Experiment 1), to be reinforced, the birds had to peck exactly four times on each key with a sequence that differed from the last trial. Thus, there were two ways for a subject to produce time-out: (a) Respond four times on the left and four times on the right with a sequence that was identical to the last sequence, or (b) respond more than four times on either of the keys. The major difference between the V and the VC conditions was that the eight responses could be distributed in any manner under the former schedule, whereas exactly four responses per key were required under the latter.

The 5 × 5 cue light matrix functioned as in Schwartz. At the start of each trial, the top left matrix light was illuminated. A peck to the left key darkened that light and lit the one to its right. Pecks to the right key darkened the presently illuminated light and lit the one immediately below it. A fifth peck to either key moved the matrix light off the board, darkening the currently illuminated light and initiating time-out. The VC procedure was continued until the percentage of reinforced trials for each subject was stable over 5 sessions, or 22 to 31 sessions.

Return to variability: Lag 5. The Lag 5 variability contingencies used in the first phase were reinstated for 6 to 20 sessions. The eight responses could again be distributed in any manner; that is, the requirement of no more than four responses per key was removed. This variability phase differed from the original in that during the first five trials of each session, the bird's sequences had to differ from the last five sequences, including sequences emitted during the last five trials from the previous session. That is, the comparison set was made continuous over sessions so that, for example, the fourth trial in a session would be reinforced only if the fourth sequence differed from the preceding three sequences of that session and the last two sequences of the previous session.

Owing to a programming error, the timing of experimental events (reinforcement, time-out, and interpeck interval) was altered when the comparison set was made continuous over sessions: All timed events were lengthened by a factor of 10/6; reinforcement and time-out were 5.83 s and interpeck interval was 0.83 s. The effect of these changes was examined in this experiment and more directly in Experiment 5, below, which shows that these timing parameter changes had no discernible effect on sequence variability.

After completion of the research with pigeons, the VIC-20 computer's random number generator was used to generate *left* and *right* responses under both V and VC conditions identical to those experienced by the pigeons. This simulation by a random generator permitted comparison of the pigeon's sequence variability with a random standard.[1]

Results

Two basic measures were derived from Schwartz (1980, 1982a). First, percentage of reinforced trials per session (number of reinforced trials divided by total trials) indicated how often the pigeons met the contingency requirements, or the probability that the sequence was correct. The second measure, which provided a more direct index of variability, was the percentage of different sequences per session (the number of distinct eight-response patterns divided by the total number of trials). A sequence was termed *distinct* if it differed in any way from all previous sequences in a given session. Note that subjects could demonstrate mastery of the contingency requirement (percentage of reinforced trials could be very high) even

---

[1] Computer-based random number generators are often referred to as *pseudorandom* or *quasi random* because the output is generated from an equation. For ease of presentation we shall refer simply to the *random generator.*

though the percentage of different sequences remained low. Under Lag 1 conditions, for example, where a sequence had to differ only from the last trial, if a bird alternated between two sequences, all trials would be reinforced. However, percentage different would be 2/50, or 4%, a low value.

Figure 1 shows the percentage of reinforced trials for each of the four subjects averaged over the last five sessions of each condition. The bars show medians for the 4 subjects' performances. (This same convention will be used in all following figures, except where otherwise stated.) Under the VC condition, shown by the middle bar, 42% of trials were reinforced per session, whereas under the Lag 5 V conditions, shown by the first and last bars, more than 90% of all trials were reinforced. (The median percentage reinforced for the 3 subjects that experienced Lag 1 V training was a similarly high 94%.)

Performances under the original and replicated V conditions did not differ statistically, and for purposes of further statistical analyses, the initial and replicated V data were averaged and then compared with VC data. The percentage of reinforced trials under V was significantly higher than the percentage of reinforced trials under VC, $t(3) = 14.44$, $p < .001$.

Figure 1 also shows that the performance of a simulated random generator (open diamonds) yielded percentage reinforcement values of 96% under the V condition and 29% under VC. In terms of successfully meeting the schedule requirements, the pigeons and random generator were similarly affected by the conditions of this experiment.

Figure 2 shows that during the last five sessions of V and VC conditions, respectively, more than 70% of response patterns differed from all previous sequences in the session. The two conditions did not differ statistically. Two pigeons showed a decrease in percentage of different sequences from initial V to VC conditions and a subsequent increase when V was reinstated; all 4 pigeons showed an increase of percentage difference scores from VC to the replicated V condition. This slight tendency for there to be a greater number of different sequences under the V condition was also observed in the simulated bird's performance, shown by the open diamonds.

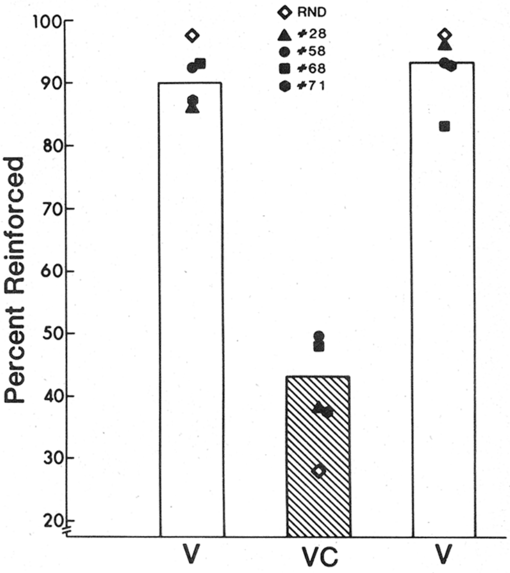

Figure 1. Percentage of reinforced trials per session during variability (V), variability-plus-constraint (VC), and replication of variability (V) conditions. (Filled points = arithmetic averages over the final five sessions; bars = medians of the pigeons' performances; open diamonds = simulated performance from a computer-based random number generator.)

Because the output of the random generator could not have been affected by the reinforcers, the only explanation is that there were fewer possible sequences under VC (where trials were terminated by a fifth response to either key) than under V (where all trials required eight responses).

Discussion

Under variability-plus-constraint conditions, only 42% of trials were reinforced, a finding that essentially replicates Schwartz (1980), in which 36% of trials ended in reinforcement. In sharp contrast, under the variability condition, 90% of trials were reinforced. The main question for discussion is why V and VC conditions yielded such different success rates. The procedural differences between the two conditions complicate interpretation: Matrix lights were present under VC and not V; sequences had to differ from the last trial under VC and from the last five trials under V; a total of 70 sequences were potentially reinforceable under VC (all sequences of eight responses having exactly

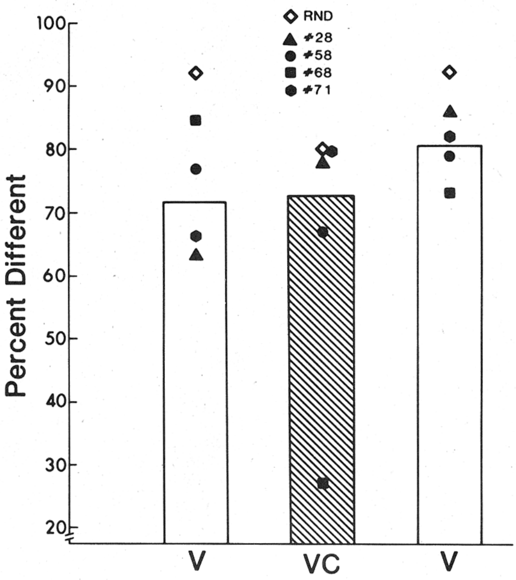

Figure 2. Percentage of different sequences (those differing from all previous sequences in a session) emitted by each subject under variability (V), variability-plus-constraint (VC), and replicated variability (V) conditions. (Filled points = arithmetic averages over the final five sessions; open diamonds = simulations from a computer-based random generator; bars = medians of the pigeons' performances.)

four responses on each key), whereas 256 sequences were reinforceable under V (all eight-response sequences); and there were two ways to commit an error under VC (repeat the last sequence or emit more than four responses on a given key), whereas only the former error was possible under V.

The fact that the simulating random generator also was almost always correct under V but only infrequently correct under VC helps to explain the obtained results. The random generator was not responding to the matrix lights. Furthermore, a random generator would be expected to gain slightly more reinforcers under a Lag 1 contingency (VC) than under a Lag 5 contingency (V). Therefore it is unlikely that either lights or lag can account for the different results under V versus VC. The remaining possible explanations involve (a) the difference in the numbers of reinforceable sequences and (b) a related factor, the different ways to make errors under the two conditions. If a pigeon responded randomly under the VC contingencies, only approximately one third of the trials would end with a reinforcer. On the other hand, the same random responses would be reinforced on more than 99% of trials under an analogous Lag 1 contingency where the constraint of no more than four responses per key was absent. An additional analysis of the computer-simulated data showed that during 250 simulated trials under the VC condition, whenever a trial was not reinforced, it was because more than four responses had been made to one key. Further analysis of the pigeons' data showed the same effect. Over the last five sessions of the VC condition, for 3 of the subjects, more than 99% of the nonreinforced trials were due to a fifth response on one of the two keys; less than 1% of the errors were due to a repetition of the last sequence. The fourth pigeon (No. 68) was a slight exception, with 81% of nonreinforced trials due to more than four responses on a key and 19% due to repetition.

Schwartz (1982a, p. 177) concluded that "reinforcement of variable response sequences in pigeons does not succeed." The present analysis suggests that the lack of success was due to the presence of the arbitrary four-responses-per-key constraint.

Experiment 2: Exact Replication of Schwartz

Parameters differed among the variability-plus-constraint and variability procedures in Experiment 1 and the Schwartz procedures. The present experiment therefore attempted to repeat the present Experiment 1 with parameters in both V and VC conditions as close as possible to Schwartz (1982a, Experiment 1). This provided a conclusive test of whether the four-response-per-key constraint was responsible for the low frequencies of reinforcement under Schwartz.

Method

Subjects and Apparatus

The subjects and apparatus were the same as in Experiment 1.

Procedure

Variability plus constraint. The procedure was identical to VC in Experiment 1 (eight responses per trial, four responses required on each key, Lag 1 look back), except that, as in the work of Schwartz (1982a), (a) there was no interpeck interval (i.e., a free-operant procedure was employed); (b) reinforcement consisted of 4 s of access to mixed grain; and (c) all trials, both reinforced and not, were followed by a 0.5-s intertrial interval. The subjects received 17 to 20 sessions, until performances were stable.

Variability. This procedure was identical to the above VC except in two respects. Trials did not terminate until the eighth peck; that is, the requirement of no more than four responses per key was omitted, as in the V condition in Experiment 1. Second, although matrix lights were used for feedback, it was necessary to reverse the direction of the lights whenever more than four pecks were emitted on a given key. For example, each of the first four pecks to the left key moved the point of matrix illumination one position to the right; each additional left-key peck moved the illumination one position back to the left. An analogous reversal occurred for the up–down direction as a function of right-key pecks. Subjects received 15 sessions under this procedure.

Results

The left portion of Figure 3 shows average percentage of reinforced trials per session over the last five sessions in each condition. These results essentially replicate Experiment 1: When pigeons were permitted to distribute their behaviors on the two keys in any fashion, they achieved reinforcement on more than 80% of the trials. However, when a four-response-per-key constraint was added, percentage reinforcements fell to slightly higher than 40%, a significant decrease, $t(3) = 6.667$, $p < 0.01$.

The right side of Figure 3 shows the percentage of sequences that differed from every previous sequence in each session averaged over the last five sessions. The percentage different scores under the VC condition were higher than under V in three of the four pigeons. That is, the birds were more variable under VC than they were under V. This difference was not statistically significant.

Figure 4 shows why the relatively high variability under VC resulted in relatively infrequent rewards. The percent of nonreinforced trials under VC (left bar) is divided into two categories, shown by the two right bars. One cause for nonreinforcement was termination of a trial by a fifth response on either key. The other cause was repetition of the immediately preceding eight-response sequence. Terminated trials accounted for 53% of all trials; repetitions accounted for less than 2%.

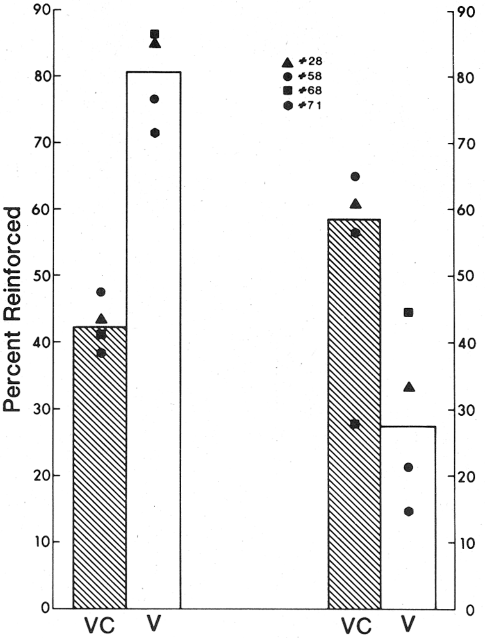

Figure 3. Percentage of reinforced trials (left two bars) and percentage of different sequences (right bars) in the variability (V) and variability-plus-constraint (VC) conditions. (Filled points = averages over the final five sessions; bars = medians of the pigeons' performances.)

Discussion

The contingencies and parameters in the present VC condition were identical to those used by Schwartz (1982a). The contingencies under V were, as far as possible, identical to the VC condition with one important exception: Under V there were no constraints on the distribution of responses. We conclude that constraining the response distribution to four responses per key in VC and in Schwartz (1980, 1982a) caused the low frequencies of obtained rewards.

The pigeons generated fewer different sequences under V than under VC. As noted above, little variability was required by a Lag 1 contingency (one successful strategy being to alternate between only two sequences). Thus, the variability engendered by the V contingencies appeared to be sensitive to the degree of variability required (see below for confirmation). Furthermore, despite the rel-

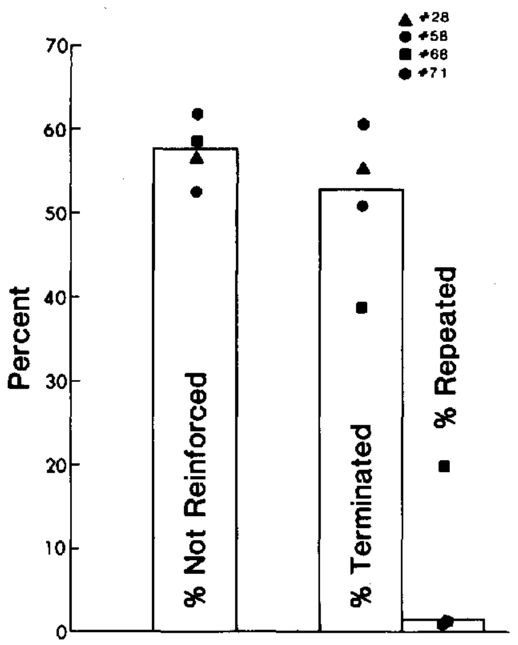

Figure 4. Percentage of trials that were not reinforced and division of nonreinforced trials into terminated (more than four responses emitted on a key) and repeated (a given eight-response sequence repeated on two consecutive trials). (Filled points = averages from final five sessions under the variability-plus-constraint condition; bars = medians of the fours pigeons' performances.)

atively high variability demonstrated by the birds in the VC condition, reinforcement frequency remained low because of the response-distribution requirement.

Although the percentage of reinforced trials under the present VC contingencies approximated those obtained by Schwartz (1980; 1982a), one result differed. Schwartz found that birds emitted few different sequences, that one modal sequence came to dominate, and that overall variability of responding was low. There are two possible reasons for the relatively high variability shown by the present subjects under VC. First, in Schwartz (1980), all pigeons were given extensive preliminary training on a sequencing task in which the only requirement was to move the light from upper left to lower right. Under this sequencing task, patterns did not have to differ from previous patterns, and all pigeons eventually came to repeat highly stereotyped patterns (e.g., LLLLRRRR). Only following this stereotypy-inducing training were the pigeons required to vary their sequences. In his second experiment (Schwartz, 1982a), birds again received prior training on the sequencing task (although the duration of this training was not stated). In the present experiments, pigeons had to respond variably (V condition in Experiment 1) before being placed under the Schwartz contingencies (VC). Second, Schwartz continued his experiments for more sessions than we did. Although there was no indication that variability under V conditions decreased with training (see Experiment 5, below, for the opposite conclusion), it is possible that with continued training under VC, response stereotypies would have developed.

Experiment 3: How Variable? Lag as Parameter

Variability is a continuum. The two previous experiments showed that pigeons earned frequent rewards under Lag 1 and Lag 5 look-back contingencies when there were no additional constraints on response distribution. This experiment asked whether pigeons could maintain high success rates when the variability requirement became increasingly stringent, that is, when the look back was increased. By the end of this experiment, to be rewarded, a pigeon had to generate a sequence that differed from every one of its last 50 trials.

Method

Subjects and Apparatus

Two experimentally naive pigeons (Nos. 70 and 73) and two previously experienced pigeons (Nos. 59 and 61) were housed and maintained as in Experiment 1. The apparatus used in this experiment was the same as in Experiment 1, but the matrix lights were not used.

Procedure

The four pigeons were autoshaped as described in Experiment 1. The two previously naive subjects (Nos. 70 and 73) then received training under Lag 1 conditions with parameters the same as in the Lag 1 variability phase of Experiment 1: Reinforcement and time-out were 3.5 s each; an interpeck interval consisting of darkened keys lasted 0.5 s; eight responses constituted a trial; the sequence on each trial had to differ from that on the last trial for reinforcement; and there were 50 trials per session. Throughout this and the remaining experiments, there were no additional response constraints; that is, the eight responses could be distributed in any manner across the two keys. After 12 or 13 sessions, the lag requirement was increased to 5: For reinforcement, sequences had to differ from those in each of the last 5 trials. The two experimentally sophisticated pigeons began the experiment at this Lag 5 value. There followed 10 to 21 sessions under Lag 10, 8 to 25 sessions under Lag 15, 10 to 38 sessions under Lag 25 (No. 59 remained at Lag 25 until the end of the experiment because its performance never reached stability), and (for the 3 remaining subjects) 23 to 45 sessions under Lag 50. Throughout the experiment, the lag value was changed only after a subject's percentage of reinforced trials had become stable over at least 5 sessions. Midway in the Lag 10 condition, the procedure was changed (as described in Experiment 1) so that the comparison trials included the final trials of the previous session. Thus, for example, under the Lag 50 condition, if the subject were responding on its 11th trial in a session, for a reward to be presented, the current sequence had to differ from the 10 trials already completed in the present session and the last 40 trials of the previous session. Because of the same programming error described in Experiment 1, from the midpoint of Lag 10 through the end of the experiment, all timed events were increased by a factor of 10/6: Reinforcement and time-out were 5.83 s rather than the original 3.5 s, and interpeck interval was 0.83 s rather than the original 0.5 s.

Results and Discussion

Figure 5 shows average percentage of reinforced trials over the last five sessions at each lag, or look-back value. The solid line connects the medians of the four pigeons, and the broken line shows a random generator simulation under identical conditions. From Lags 5 through 25, more than 85% of the pigeons' sequences met the variability requirements and were reinforced. At Lag 50, there was a decrease to 67%. This same decrease in percentage reinforced was seen in the random generator's simulated data. Thus, the pigeon's data again paralleled the data of a random generator.

To obtain the high frequencies of reinforcement shown in Figure 5, the pigeons must have generated highly variable response sequences. One index of this high variability is shown in Figure 6. As the lag requirement increased from 5 to 25, the percentage of sequences that differed from all previously emitted sequences in a session increased from 66% to 87%. As discussed above, to maintain a high frequency of reinforcement under a low lag requirement, the bird did not have to emit very many different sequences. As the length of the look back increased, however, increasingly greater numbers of different sequences were demanded. Sensitivity to these changing requirements is indicated by the increasing function from Lag 5 to Lag 25. The small decrease to 81% different at the

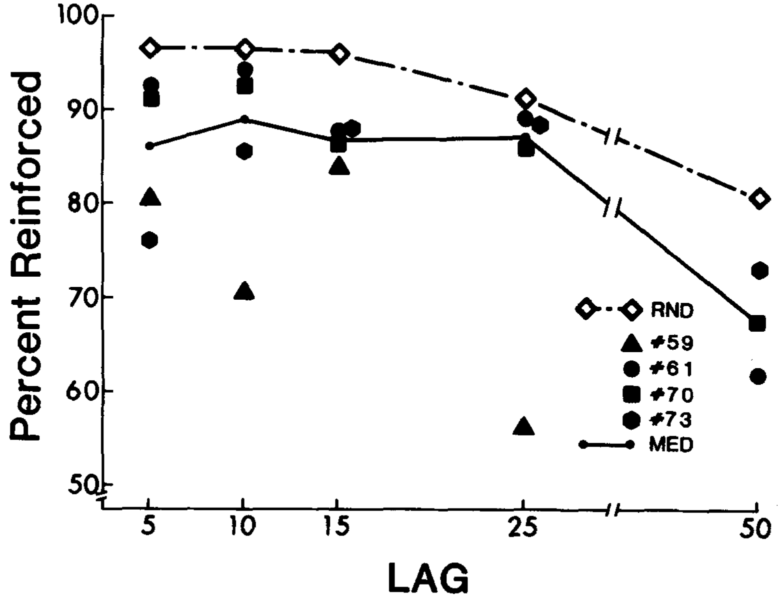

Figure 5. Percentage of reinforced trials per session as a function of lag (look back) value. (Filled points = averages over final five sessions at each lag value for each subject; solid line = medians; broken line = performance of a simulating random generator under identical conditions.)

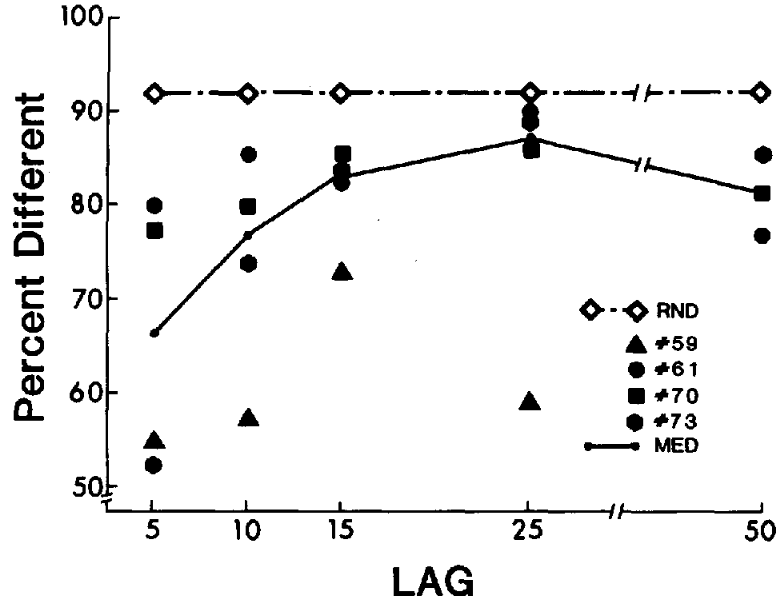

Figure 6. Percentage of different sequences per session as a function of lag value. (Filled points = averages over final five sessions at each lag value for each subject; solid line = medians; broken line = performance of a simulating random generator under identical conditions.

Lag 50 value may be a respondent effect, correlated with the lowered frequencies of obtained reinforcements. As shown by the performance of the random generator in Figure 5, no matter how variable the behavior, as lag requirement continued to increase, a decrease in the number of rewards gained resulted.

More detailed indices of overall response variability are given in Figure 7, which shows three measures of variability (U values, or average uncertainty) as functions of lag. The U values were computed according to the following equations:

$$U_1 = \frac{\sum_1^2 p_i \log_2 p_i}{\log_2(2)},$$

$$U_2 = \frac{\sum_1^4 (p_i \log_2 p_i) - U_1}{\log_2(4)},$$

and

$$U_3 = \frac{\sum_1^8 (p_i \log_2 p_i) - U_2}{\log_2(8)},$$

where, for U_1, p_i equals the probabilities of L and R responses; for U_2, p_i equals the probabilities of LL, LR, RL, and RR response pairs; and for U_3, p_i equals the probabilities of LLL, LLR, LRL, LRR, RLL, RLR, RRL, and RRR triplets. The U measure, derived from information theory (Miller & Frick, 1949), varies between 0.0 and 1.0, with 0.0 indicating that all responses are perfectly predictable or ordered and 1.0 indicating maximum uncertainty. The U values were calculated by concatenating all responses without regard to reinforcement or time-out and computing the relative frequencies of left and right responses (U_1), pairs of responses (U_2), and triplets of responses (U_3). When left and right were approximately equal, U_1 approached 1.0; when all possible pairs of responses approached equality, U_2 approached 1.0; and when all possible sequences of responses taken three at a time approached equality, U_3 approached 1.0. Figure 7 shows that as lag values increased (i.e., as requirement for variability increased) the averages of the 4 pigeons' U_1, U_2, and U_3 values increased. (These averages well represent the individual functions and are presented to save space.) At the lag value of 25, the average of the 4 pigeons' U values approximated the U value of the random number generator (shown by the open diamonds; only a single line is drawn because U_1, U_2 and U_3 values were approximately the same for the random generator).

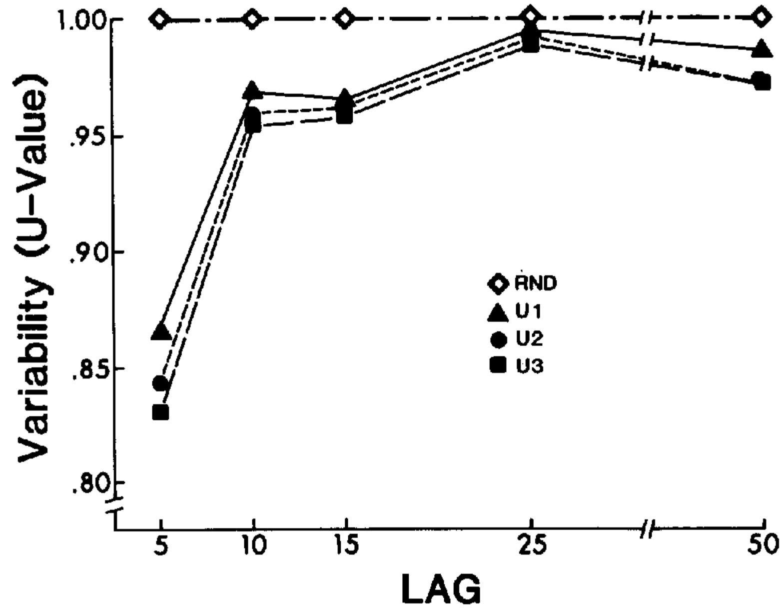

Figure 7. Average sequence variability as functions of lag. (Each of the lines connects the medians of the pigeons' average performances over final five sessions of each lag value. U_1 = uncertainty for responses taken one at a time; U_2 = uncertainty for responses considered in pairs; U_3 = uncertainty for response triplets. Open diamonds = analogous data for simulating random number generator, where, because the three U values were almost identical, a single point indicates the values.)

Once again, there was a slight tendency for variability to decrease at Lag 50. The closeness of the three U functions to one another indicates high variability and absence of higher order biases or stereotypies. (Note that we also examined U_4 through U_8, and these contained similar information to that shown in U_1 through U_3.

Two possibly confounding influences were present in Experiment 3. First, the timing parameter changed during the Lag 10 condition, and therefore the Lag 15 through Lag 50 conditions contained longer reinforcement and time-out durations than the Lag 5 and Lag 10 conditions. However, Experiment 1 showed that there was no statistically significant effect of these timing differences, a result supported in Experiment 5, below. Second, the lag requirements increased from low values to high, and therefore the form of the obtained function may partly be due to the order of experience. We thought that pigeons could not tolerate high lag requirements before they experienced training under lower requirements, a hypothesis shown to be incorrect in Experiment 5. The general form of the subjects' percentage reinforced function paralleled that of the simulated random function; this finding was again consistent with the hypothesis that the pigeons were generating quasi-random sequences. Although the variability of the computer-based random generator was, of course, unaffected by the reinforcement schedule, the pigeons' variability appeared to be controlled by the reinforcers. When the schedule demanded relatively little variability (Lag 5), variability was relatively low. As the variability requirement increased to Lag 25, so too did variability of performance. However, at Lag 50, when the obtained frequency of reinforcement decreased despite random performance (as indicated by the simulating random generator), again the birds' variability decreased. Thus, high variability was engendered only when it was differentially reinforced.

Experiment 4: Quasi-Random Versus Memory Strategies: Number of Responses as Parameter

The present experiment asked how the pigeons generated their variable response sequences: What mechanism or strategy accounts for such highly variable performance? The previous experiments alluded to one possible strategy, that of a quasi-random generator, but alternative strategies involving remembering previous sequences would also do. For example, a subject could learn a long random response sequence (see Popper, 1968) or utilize a rational system (e.g., first emit 8 left responses; then 7 left and 1 right; then 6 left, 1 right, and 1 left; etc.). It is not here being suggested that pigeons can count in binary but rather that their behavior could be described as systematic. Any systematic strategy would involve a rather large memory load, for the pigeon would have to remember where it was in its system. The present experiment attempted to contrast the quasi-random and memory hypotheses in the following way. If the number of responses required for each sequence were increased, performance based on a memory strategy should be adversely affected, for it is easier to remember a four-response sequence than an eight-response sequence. On the other hand, if the bird were acting as a quasi-random generator, success rates should improve with increasing responses per trial, because by the laws of chance, a random generator would be more likely to repeat sequences comprising four responses (1 chance in 16 under Lag 1) than eight responses (1 chance in 256). Thus, the memory and quasi-random generator hypotheses make opposite predictions. If the subject's rewards increased with increasing responses per trial, the quasi-random hypothesis would be supported; if the rewards decreased with increasing responses per trial, a memory strategy would be supported.

Method

Subjects and Apparatus

The subjects and apparatus were the same as in Experiment 1.

Procedure

The procedure was identical to that in Experiment 3, except the lag requirement was kept constant at Lag 5 and the number of responses per trial, or sequence length, varied in an ABCB format. The pigeons were first given 24 to 29 sessions under a six-responses-per-trial condition (there were 64 possible sequences); then 12 to 38 sessions under eight responses per trial (identical to the Lag 5 variability conditions in Experiments 1 and 3 where there were 256 possible sequences); 9 to 23 sessions under four responses per trial (16 possible sequences); and another 6 to 9 sessions under eight responses per trial. The reinforcement and time-out intervals were 5.83 s throughout, and interpeck interval was 0.83 s.

Results and Discussion

Figure 8 shows the average percentage of the 50 available rewards obtained over the last five sessions at each responses-per-trial value. The eight-response value represents the mean of the two eight-response phases, which were statistically indistinguishable from one another. For all subjects, as number of responses per sequence increased, percentage of reinforced trials increased monotonically. An analysis of variance with repeated measures showed an overall significant difference among the three conditions, $F(2, 6) = 38.21$, $p < .001$, and analytical pairwise comparisons showed that each condition differed from every other: four versus eight responses, $F($1,

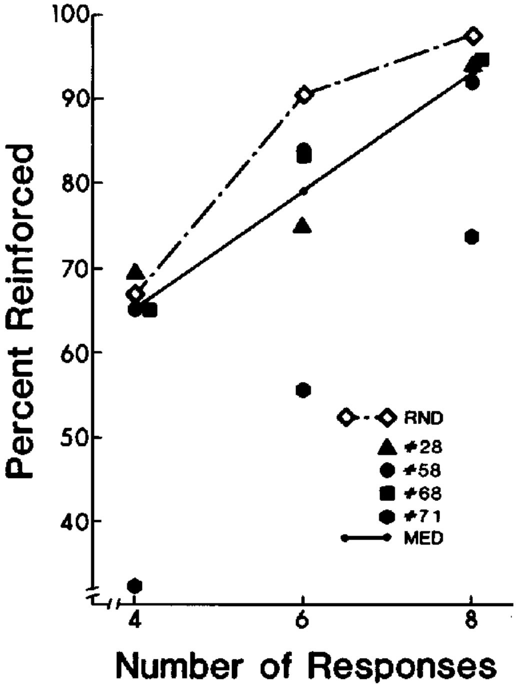

Figure 8. Percentage of reinforced trials as a function of number of responses per trial. (Filled points = averages over final five sessions at each condition for each of the pigeons; solid line = medians; broken line = data from simulating random generator under identical conditions.)

6) = 77.37, $p < .001$; six versus eight responses, $F(1, 6) = 16.17$, $p < .01$; and four versus six responses, $F(1, 6) = 22.80$, $p < .005$. The random number simulator's percentage of reinforced trials, shown by the open diamonds, also increased monotonically as a function of increasing responses per trial.

Figure 9 shows that U values (measures of overall response variability, as explained in Experiment 3) varied nonmonotonically over a relatively small range. (Note that the percentage different statistic is inappropriate in the present case because the number of possibly different sequences varied with responses per sequence.) Analyses of variance showed no significant effects for U_1, U_2, or U_3. Thus, once again, the pigeons' function approximated the simulated random generator (open diamonds), and the quasi-random hypothesis was supported.

Experiment 5: Is Variability a Reinforceable Dimension? Lag 50 Versus Yoked Variable Ratio

Neither the experiments described above nor any previously published study conclusively demonstrated that variability is an operant dimension controlled by reinforcement. A distinction was drawn at the beginning of this article between elicited, or respondent, effects of reinforcement schedules and reinforcing effects. For example, in Herrnstein (1961), variability of responding along a strip was monitored as a function of different reinforcement schedules. As long as the pigeon pecked anywhere along the strip, responses were effective. Variability was therefore orthogonal to those dimensions required by the schedule contingencies; reinforcement did not depend on variability. (In Herrnstein and the other respondent cases described at the beginning of the article, variability may have been reinforced adventitiously. See Neuringer, 1970, for a related analysis of superstitious key pecking in pigeons.)

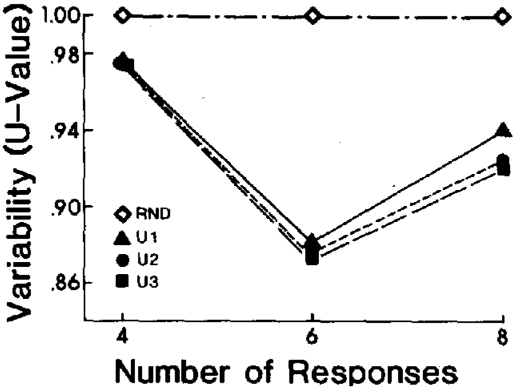

Figure 9. Average response variability as a function of number of responses per trial. (U_1 = uncertainty measured by taking responses singly; U_2 = uncertainty by taking pairs of responses; U_3 = uncertainty by taking response triplets; open diamonds = uncertainty of responses generated by simulating random generator.)

The question now raised is whether the variability observed in Experiments 1 through 4 was a by-product of the particular schedules used or whether the observed variability depended on direct reinforcement of that variability. For an answer, subjects were first presented with a schedule in which, to be reinforced, each eight-response sequence had to differ from the previous 50 sequences (Lag 50 variability condition). After stable performances were attained, each pigeon was presented with the exact same frequency and pattern of rewards that it had received over its last six sessions of Lag 50 variability, but now the rewards depended only on an emission of eight responses and not on sequence variability. With this self-yoking procedure, we determined whether the variability observed under a Lag 50 schedule was due to respondent effects of the schedule or to reinforcement of operant variability. A reinforcement-of-variability hypothesis would be supported only if sequence variability were appreciably higher under the Lag 50 variability condition than under the self-yoke condition.

Method

Subjects

Two experimentally naive (Nos. 49 and 50) and two previously experienced (Nos. 44 and 45) pigeons were maintained as described in Experiment 1.

Apparatus

Two 30-cm × 40-cm × 30-cm chambers made of aluminum walls, with a Plexiglas door inset in the rear and a wire mesh floor and ceiling were housed in two sound-attenuating outer boxes. A Gerbrands masking noise generator provided additional sound masking. On the front wall of each chamber were three 2-cm-in-diameter Gerbrands response keys with their centers 7.5 cm from each other and from the side walls and 21.5 cm above the mesh floor. Keys could be transilluminated with 7.5-W blue bulbs. A response registered when applied force was withdrawn. The middle of the three keys was covered with black tape and could not be operated. Directly below this middle key was a round hopper opening, 5 cm in diameter, with its midpoint 10 cm from the floor, through which a Gerbrands magazine could provide mixed pigeon grain reinforcement illuminated with a 7.5-W white bulb. House light was provided by two 7.5-W white bulbs above the wire mesh ceiling. As described in Experiment 1, VIC-20 computers controlled the experiment. Each pigeon was arbitrarily assigned to one of the two experimental chambers.

Procedure

The four subjects were first given autoshaping training as described in Experiment 1. The experimental procedure then followed an ABAA'B design, with A and A' representing Lag 50 variability contingencies and B representing a yoked variable ratio (yoked-VR) contingency in which variable sequences were not required. The variability procedure was identical to that in Experiment 1 except as follows. A Lag 50 requirement was present from the outset. As in the latter phases of Experiment 1, the comparison criterion was continuous over sessions. For example, for Trial 10 to be reinforced, the sequence of responses in that trial had to differ from each of the sequences in the 9 previous trials of the current session and the last 41 completed trials of the preceding session. The initial Lag 50 phase continued for 26 to 38 sessions or until each pigeon maintained a stable percentage of reinforced trials for 5 or more sessions. Throughout the present experiment, sessions terminated after the 50th reinforcer, or after 100 trials, whichever occurred first.

At the start of the B phase, the contingencies of reinforcement were changed so that the pigeons were reinforced on a yoked-variable ratio schedule derived from their individual performances under Lag 50. Under this yoked-VR, eight responses again constituted a trial, and trials were again sometimes followed by grain reinforcers and sometimes by time-out, but reinforcement and time-out presentation were now independent of sequence variability. Each pigeon's last 6 sessions under the Lag 50 variability contingencies were used to create its schedule of reinforcement under yoked-VR. Thus, each subject was yoked to itself, and the schedule of reinforced and nonreinforced trials under yoked-VR replicated the pattern of reinforcers and time-outs obtained under Lag 50 variability. The yoked reinforcement schedule lasted for 6 consecutive sessions and then was repeated. To illustrate, if Subject 44 had been rewarded after Trials 2, 5, 6, and 8, and so on, in the last session under the Lag 50 condition, then in yoked-VR Sessions 6, 12, 18, and so on, Subject 44 would be rewarded after Trials 2, 5, 6, and 8, and so on, regardless of which eight-response sequence was emitted. Trials 1, 3, 4, and 7, and so on, would be terminated by time-out. The yoked-VR contingencies continued until performance was stable, from 24 to 31 sessions, whereupon Lag 50 variability contingencies were reinstated and maintained for 17 or 18 sessions.

Due to the programming error described in Experiment 1, the duration of events (reinforcement, time-out, and interpeck interval) was shorter (by a factor of 10/6) in yoked-VR than in either the preceding or the following Lag 50 phases. The reinforcer and time-out were 3.5 s in yoked-VR, as opposed to 5.83 s in Lag 50. The interpeck interval was 0.83 s in Lag 50 phases but 0.5 s in the yoked-VR. To compare directly the effects of the different time values on sequence variability, after each pigeon had reached stability on Lag 50 (second A phase containing the long times), the durations of the reinforcer, time-out, and interpeck interval were changed to those under yoked-VR with everything else held constant, thus permitting comparison of responding under Lag 50 long times with Lag 50 short times. After stable performances were reached in 10 to 20 sessions, the yoked-VR contingencies were reinstated for 17 to 32 sessions, thereby permitting comparison of Lag 50 and yoked-VR when all time parameters were identical. Schedules of reinforcers and time-outs in this second yoked-VR phase were derived from performances of each pigeon in the last 6 sessions of the A' Lag 50 (short times) phase.

Results and Discussion

The main results were that (a) variability was significantly higher under Lag 50 than yoked-VR, thereby demonstrating that the variability depended on contingent reinforcement; (b) experimentally naive pigeons placed directly on the Lag 50 schedule very quickly learned to vary, thereby indicating that variability was easy to condition; and (c) the 10/6 differences in reinforcer, time-out, and interpeck interval times had no discernible effect on performance, thereby indicating the robustness of the reinforcement-of-variability effect.

The left two bars of Figure 10 show mean percentages of different sequences per session for the first five (open bar) and last five (striped bar) sessions under the first Lag 50 (long times) schedule. Over the first five sessions, more than 50% of sequences differed from all previous sequences in the same session, and this value increased to more than 75% by the last five sessions. By the end of this Lag 50 phase, the pigeons were being rewarded after approximately 70% of their trials, an increase from 50% during the first five sessions. The increases from first to last five sessions in both percentage different trials, $t(3) = 6.68$, $p < .01$, and percentage reinforced trials, $t(3) = 5.11$, $p < .025$, were statistically significant.

The second set of bars, representing yoked-VR, shows that when the variability contingencies were removed, percentage of different sequences fell immediately until fewer than 20% of the sequences were different from previous sequences in the session. The difference between the last five sessions of the Lag 50 variability contingencies and the last five sessions of the yoked-VR was significant, $t(3) = 7.46, p < .005$.

Upon reintroduction of the Lag 50 contingencies, percentages of different sequences rose immediately. The leftmost of the four bars above the repeated Lag 50 condition shows percentage different during the first five sessions of Lag 50 following yoked-VR, and the second of the four bars shows the last five sessions under this phase. Because of the differences in timing values (long times in Lag 50 and short times in yoked-VR), the observed effects might have been confounded. The reinforcement, time-out, and interpeck interval times were therefore changed under Lag 50 so that these times were now identical to those under yoked-VR. The third and fourth bars above the replicated Lag 50 show percentage different sequences during the first five and last five sessions under short time values. There was essentially no change in percentage of different sequences due to the different time values.

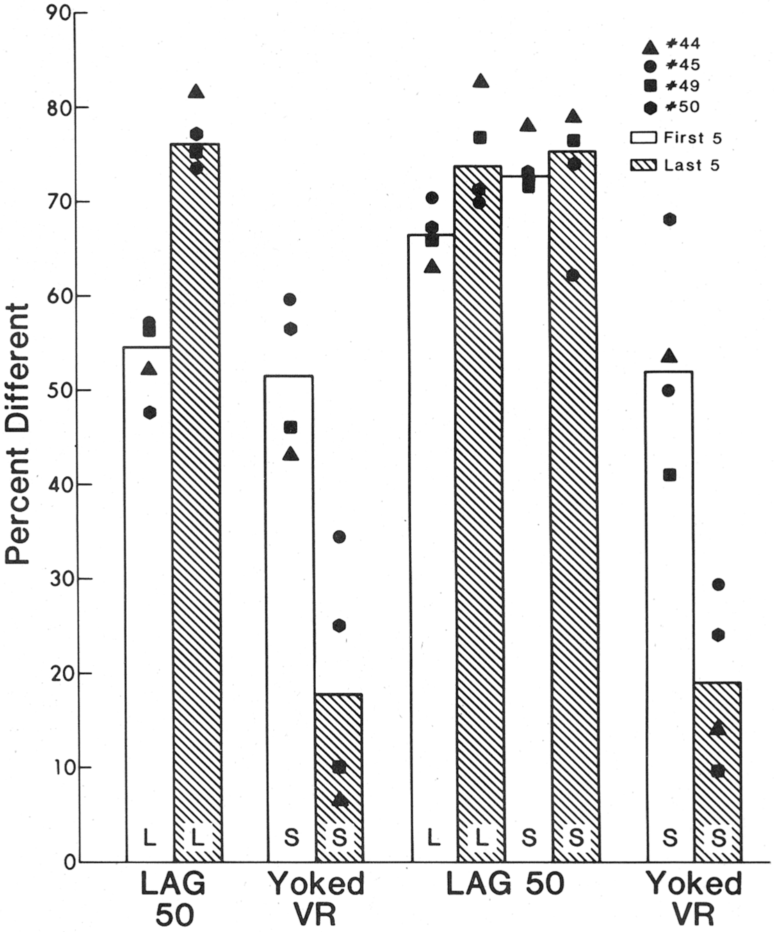

Figure 10. Percentage of different sequences per session under one condition where each sequence had to differ from the previous 50 sequences for reinforcement (Lag 50) and another condition where reinforcements were given independently of response variability (yoked VR). (Open bars = first five sessions of each condition; striped bars = final five sessions. L = long timing values; S = short times. Each point = 1 pigeon's arithmetic average performance over five sessions; bars = medians of the 4 pigeons' performances.)

Upon return to the yoked-VR condition, variability of responding again decreased immediately, and once again the difference between the last five sessions of Lag 50 (now with short times) and last five sessions of yoked-VR (with identical short times) was significant, $t(3) = 6.771$, $p < .01$.

The fact that response variability depended on contingent reinforcement is shown in Figure 11 by the percentage of modal sequences, defined as the single sequence emitted more frequently than any other in a session. The ordinate shows the percentage of trials per session in which the modal sequence occurred. By the last five sessions of the first phase of Lag 50, the modal sequences accounted for about 4% of the sequences emitted per session. On the other hand, by the end of the first yoked-VR phase, modal sequences accounted for almost 50% of the sequences. Absence of significance between Lag 50 and yoked-VR, $t(3) = 2.493$, $p = .0873$, was due to the large spread of the individual subjects' data under the yoked-VR condition. When the schedule demanded high variability (Lag 50), all subjects emitted very few repetitions of any given sequence; when the schedule permitted variability but did not require it (yoked-VR), there were large intersubject differences. These same patterns of modal frequencies were replicated with return to Lag 50 (only the short time phase of Lag 50 is shown in the figure) and then to yoked-VR, with the difference during the last five sessions of these conditions being statistically significant, $t(3) = 3.46$, $p < .05$. In almost all cases, the sequence defined as modal under yoked-VR represented exclusive responding on one or the other key (e.g., eight left responses or eight right responses). Because reinforcement did not depend on any particular sequence, the final behavior was

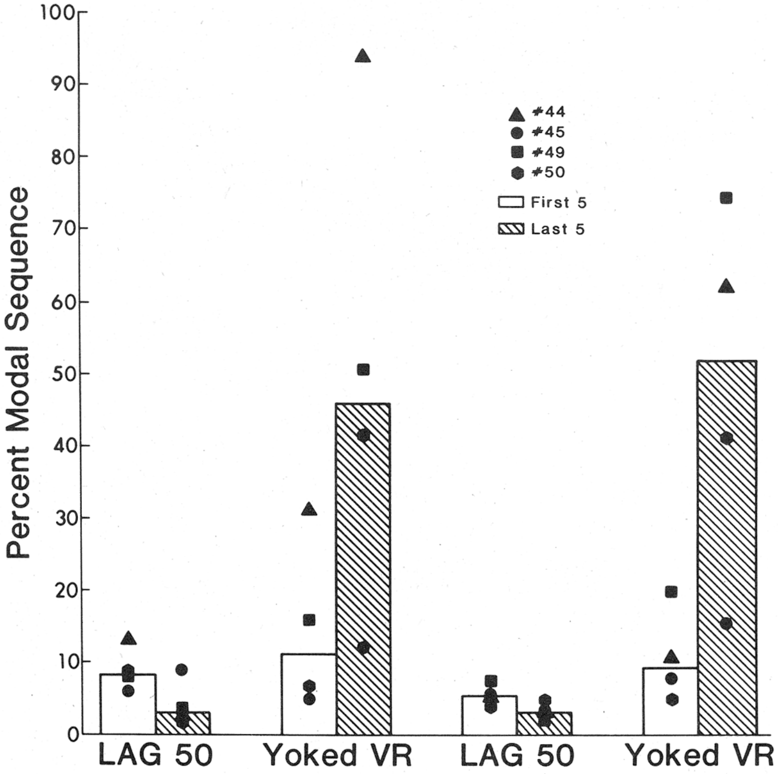

Figure 11. Percentage of modal sequences per session (number of trials in which the most common pattern occurred divided by the total number of trials) as a function of Lag 50 versus yoked variable ratio (yoked-VR) conditions.

probably a function of minimizing energy expenditure (it takes more energy to alternate keys than to respond on a single key) and adventitious reinforcement of superstitious patterns. Under Lag 50, however, the modal sequences generally comprised a mixture of responses on the two keys (e.g., LLRLRLLL, RLLLRRRR, LRLLRRRR, and RRRRRLRR [one example from each of the 4 pigeons]).

Average U values, an index of overall response variability as discussed above, are plotted for each of the conditions in Figure 12. For comparison, the random number generator's U values under identical simulated conditions are also drawn. During the last five sessions of both Lag 50 conditions, the pigeons' U values closely approximated the U value of the random generator. Under both yoked-VR conditions, however, the pigeons' U values were greatly lowered. The relatively low U_1 values show that the birds were forming position preferences. The differences between U_1 and U_2 and, similarly, between U_1 and U_3 show that in addition to the position preferences, second- and third-order patterns of responding were being generated with high probability.

Whether the results are considered in terms of percentages of different sequences, relative frequencies of modal sequences, or average uncertainty across the entire array of responses, a single robust conclusion is reached: The variability-requiring Lag 50 condition caused significantly more variability of sequence patterns than did the yoked-VR. The behavioral variability generated in the present experiment depended on the variability's being reinforced. Absence of high variability under yoked-VR may indicate that variable responding is an energy-expensive strategy and, hence, nonpreferred. That there was some variability under the yoked-VR condition may indicate a small respondent effect.

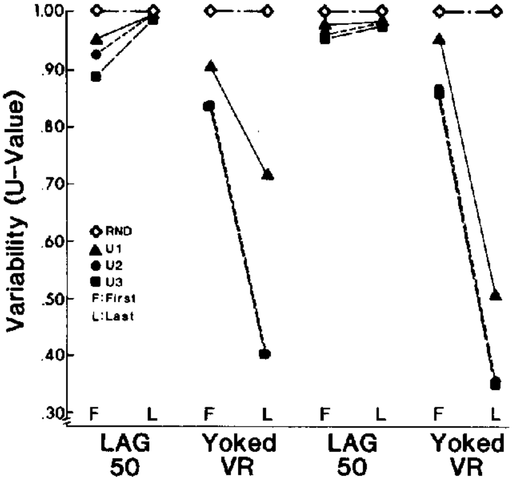

Figure 12. Average response variability under Lag 50 and yoked variable ratio (yoked-VR) conditions. (U_1 = responses taken one at a time; U_2 = responses in pairs; U_3 = responses in triplets. Open diamonds are from a simulating random generator under identical conditions. F = averages over the first five sessions; L = averages over the final five sessions of each condition.)

Experiment 6: Stimulus Control: Multiple Variability–Stereotypy

Operant, or reinforceable, dimensions of behavior (location, duration, rate, force, and topography) are sensitive to stimulus control. The question raised in this final experiment is whether the variability dimension can also come under stimulus control. In particular, can pigeons learn to generate variable sequences in the presence of key lights of one color and stereotyped, or fixed, sequences in the presence of a different color? An affirmative answer would support the thesis that behavioral variability is controlled by environmental contingencies.

Method

Subjects and Apparatus

Four experimentally naive pigeons (Nos. 29, 31, 38, and 39) were maintained as described in Experiment 1. For the apparatus, the two chambers described in Experiment 5 were again used. The key lights could be illuminated with either blue or red 7.5-W bulbs.

Procedure

An autoshaping procedure was identical to that in Experiment 1, with one exception. A trial consisted of one of four equally probable events: left key light red, left key light blue, right key light red, and right key light blue. Subjects received 5 sessions of autoshape training after their first response. Because Subject 38 did not peck the key after 20 sessions of autoshaping, the experimenter shaped key pecking by reinforcing successive approximations to the key peck response and then provided 5 additional autoshaping sessions.

There were three experimental phases. The first examined acquisition of stimulus control over variable and stereotyped responding. Following an informal exploration of the effects of parameter manipulations, the second phase attempted to equalize the number of responses per trial in the two conditions and to generate approximately equal and intermediate levels of percentage of reinforced trials. A reversal of stimulus conditions followed.

Phase 1: Acquisition of stimulus control. Following autoshaping, the four pigeons were put on a multiple variability–stereotypy schedule in which the two components alternated after every 10th reinforcement. In the variability component, both key lights were blue, and subjects had to meet a Lag 5 variability criterion identical to that described in Experiment 1, where reinforcement was contingent on a sequence of eight pecks on the two keys that differed from each of the preceding five sequences. The Lag 5 criterion was continuous with respect to other V components in the same session and across sessions.

In the stereotypy (S) component, both key lights were red and the pigeons had to emit an arbitrarily defined three-response sequence, namely, LRR, in that order. The first two correct responses in the sequence produced the same interpeck interval as in the V component, and the third correct response produced the reinforcer. An error during an S trial (e.g., if the second response were to the left key) immediately produced the same time-out as in V and reset the sequence to the beginning. Whereas the time-out in V occurred only after the eighth response, the time-out in S immediately followed any incorrect response in the sequence.

The V and S components continued to alternate after every 10th reinforcement until the bird earned a total of 60 rewards or until a maximum of 120 trials was reached in either component, whichever occurred first. After initial longer times, the durations of reinforcement and time-out were reduced to 5.0 s in both components (times for Subject 38 were decreased to 4.2 s after five sessions owing to its weight gain). Pecks in both V and S were separated by a 0.83-s interpeck interval.

Because most of the pigeons were at first unable to attain 10 rewards in S, the first few sessions began with the V component. After the 8th session, when responding in S had improved, the V and S components alternately began each session. In an attempt to further improve performances during S, the salience of time-out was increased by flashing the house light on for 0.33 s and off for 0.08 s during time-out in both the V and S components. Phase 1 training continued for 12, 16, 24, and 24 sessions for each of the 4 pigeons, respectively.

Phase 2: Equalization of responses. An attempt was made to equalize the number of responses required in V and S components and to approximately equalize the percentages of correct responses at an intermediate level of proficiency to avoid ceiling and floor effects. At the end of this phase, which lasted approximately 50 sessions, the schedule in the V component was six responses, Lag 10. The schedule requirement in S was the fixed pattern LRRLL. Reinforcement and time-out in both components lasted for 3 s.

Phase 3: Stimulus reversal. At the start of Phase 3, the number of responses in V was reduced to five to equal the number of responses in S. Otherwise, the contingencies in effect at the end of Phase 2 were maintained. Birds 29, 31, 38, and 39 received 10, 14, 12, and 22 sessions, respectively.

The key light colors signaling V and S were then reversed, with red key lights now associated with V and blue key lights with S. No other changes were made. Two pigeons (Nos. 29 and 38) began the first session of reversed color in the S component, and the other two began in V. There were 18 to 24 sessions of reversal.

Results

Figure 13 shows percentages of reinforced trials in variability and stereotypy components separately for 1 subject (No. 31) during each session of Phases 1 and 3. The performance shown is representative of all birds. During acquisition of performance under the first multiple schedule, shown in the left panel, all birds initially received a higher percentage of reinforcements in the variability component (eight responses, Lag 5) than in the stereotypy component (three responses). Over sessions, performances in both components improved in accuracy, indicating that all subjects acquired the variability–stereotypy discrimination.

Figure 14 shows average *U*-value measures—U_1 showing variability for responses taken one at a time, U_2 for responses taken two at a time, and U_3 for responses taken three at a time—in each of the components. The left set of bars indicates the median performance over the last five sessions in the initial condition. As would be expected if the pigeons were performing variably in V and emitting a fixed sequence in S, there were large and significant differences between the *U* values in the two components: For U_1, $t(3) = 15.515$, $p < .001$; for U_2, $t(3) = 15.617$, $p < .001$; and for U_3, $t(3) = 11.533$, $p < .005$.

During the exploratory phase that is not shown, attempts were made to increase the stereotypy component to eight responses (equal to the variability component) and to change the contingencies in S so that time-outs for incorrect sequences occurred only at the end of the sequence (as was the case in V). But both of these attempts failed. Therefore, to equalize performances under V and S, a five-response sequence was used in each component and the lag criterion was increased to 10 in the V component. Performances under these conditions are shown in the middle panels of Figures 13 and 14. When the stimulus conditions were reversed (blue key lights now signifying the S component

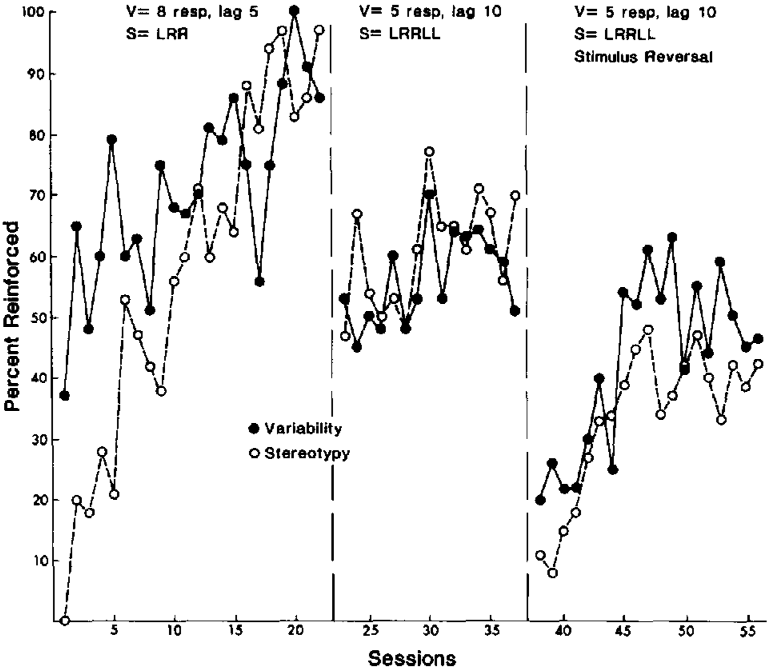

Figure 13. Percentage of reinforced trials per session in variability (V) and stereotypy (S) components of the multiple schedule for Pigeon 31. (Left panel = initial acquisition where the contingencies in V were eight responses, Lag 5, and the S contingencies reinforced left–right–right patterns; middle panel = performance under a V schedule of five responses, Lag 10, and an S schedule of left–right–right–left–left; right panel = reversal of the middle schedule conditions [key light colors were reversed].)

and red key lights signifying the V component) performances immediately deteriorated but then improved (right panels). At the end of both these phases, *U* values in S and V differed significantly at the .01 level or better. Thus, we conclude that stimulus control was established over variable and stereotyped responding.

General Discussion

Experiments 1 and 2 showed that when hungry pigeons were given grain for generating sequences of eight responses that differed from their last sequence, they successfully varied their sequences. When, in addition to meeting this same variability requirement, the pigeons had to peck exactly four times on each key (as in Schwartz, 1980, Experiment 4; 1982a, Experiment 1), success rates fell significantly. We concluded that the inability of pigeons to gain high rates of reward under the Schwartz procedure was due to an artifact of the four-response-per-key requirement. This conclusion was strengthened by comparison of the pigeon's performance with a computer-based random number simulation.

Experiment 3 increased the look back, or the number of prior sequences of eight responses from which the current sequence had to differ. Eventually, to be reinforced, the pigeon had to respond with a sequence that differed from each of its last 50 sequences. This look back included sequences from the previous session. The subjects generated highly variable sequences, with more than 80% of the patterns differing from all previous patterns in a session. Probabilities of correct sequences again paralleled the probability of correct sequences of the simulating random generator.

Experiment 4 compared two possible accounts of this variability. The memory hypothesis was that the pigeons learned a long sequence of responses or used a rational strategy to meet the schedule requirements. The variability hypothesis was that the pigeons behaved as a quasi-random generator. The first hypothesis predicts that increasing the number of responses per trial would be correlated with a lowered success rate, for it is easier to remember fewer responses than more. The quasi-random hypothesis made the opposite prediction: By chance, given few responses per trial, consecutive trials would repeat one another; given many responses per trial, there would be few repetitions by chance. When the required number of responses per trial was increased from four to eight, success rates improved significantly, thereby supporting the quasi-random generator hypothesis. Once again, the pigeons' performances paralleled the performance of a simulating random generator.

The first four experiments generated high behavioral variability. However, neither these nor any previously published experiments demonstrated that response variability depended on the contingency between variability and reinforcement. The observed variability

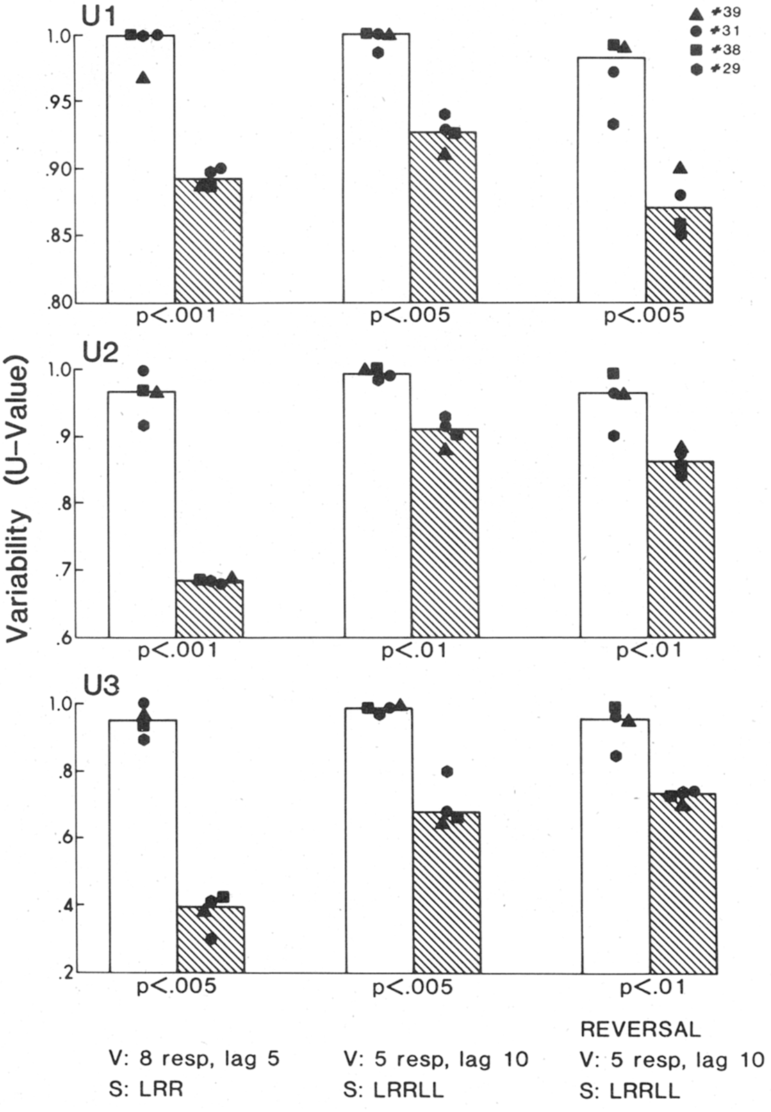

Figure 14. Average response variability under three phases of Experiment 6, from left to right. (U_1 = responses taken one at a time; U_2 = responses taken in pairs; U_3 = responses taken in triplets.)

could have been elicited by the reinforcement schedule (a respondent effect) rather than directly reinforced (an operant effect). Experiment 5 tested these alternatives by comparing the performance of pigeons under two identical schedules, except that one schedule required variability whereas the other permitted it. In the first, a pigeon had to respond eight times with a sequence that differed from each of its last 50 sequences. The patterns and frequencies of rewards under this condition were duplicated to form a self-yoked schedule where eight responses were required for reward but sequence variability was no longer necessary. The results showed that sequence variability was generated only when it was required. Under the yoked schedule, variability decreased significantly and the pigeons responded with highly repetitive patterns. Variability is therefore an operant dimension of behavior.

Experiment 6 demonstrated discriminative control over behavioral variability. Pigeons learned to respond variably in the presence of key lights of one color and with a fixed, stereotyped sequence in the presence of a second color. When the stimulus conditions were reversed, performances reversed. Thus, the variability dimension of behavior is controlled by environmental stimuli in much the same manner as other operant dimensions are.

The present series of experiments therefore conclusively demonstrate the existence and strength of operant variability, variability that is engendered and maintained because presentation of a reinforcer depends on the variability. This conclusion is consistent with Blough (1966), Pryor et al. (1969), and Bryant and Church (1974), among others.

Previous studies have examined respondent variability. Different schedules of reinforcement reliably engender differing degrees of behavioral variability with no contingency between the variability and the reinforcement schedules. For example, if identical and independent fixed-ratio 5 (FR 5) schedules are programmed on each of two response keys, most pigeons peck exclusively on one or the other key. If the schedules are changed to FR 150, there is considerable switching between keys. This observation from our laboratory is a clear example of respondent variability caused by reinforcement schedules. Variability was neither required nor differentially reinforced. The state of the environment, as well as the contingencies in the environment, influence behavioral variability, the former through respondent effects and the latter through operant effects.

Both respondent and operant variability may be adaptive. When reinforcers are infrequent or absent, variability increases the likelihood that the animal will improve its lot—learn a new strategy for obtaining reinforcement or change its environment. Even when reinforcement densities are relatively high, variability may improve the schedule or provide knowledge of the environment in anticipation of possible future decrements in reinforcement. Variability is an adaptive response to a changing or potentially changing environment.

Operant variability has unique adaptive functions not shared by respondent variability. Whenever an animal is operantly conditioned to generate a new response, whether the conditioning is through the process of shaping (successive approximations to some desired goal response) or trial and error (Thorndikian conditioning), it is adaptive for the animal to vary its behaviors. Reinforcement works through selection of some previously emitted behavior. If the to-be-selected behavior does not occur, reinforcement cannot select. On the other hand, in an environment where an operant response has previously been learned and is now being maintained by an acceptable schedule of reinforcement, high variability may not be functional, for variation may require relative high energy expenditure and sometimes result in less frequent reinforcement. It is advantageous for an animal to discriminate situations in which new responses must be learned from those in which previously learned behaviors must be repeated. We hypothesize that this discrimination is based on the reinforcement of diverse responses and response classes in the former case versus reinforcement of fixed, or stereotyped, responses and response classes in the latter. (The discrimination is in some ways analogous to that between contingent and noncontingent reinforcement; Killeen, 1978). When an animal is differentially rewarded for a variety of responses, it generates variable behaviors. We posit that this describes all operant learning (as opposed to operant maintaining) situations. Explicit reinforcement of variable behaviors prior to initiation of operant learning, or shaping, procedures might speed the learning process. This hypothesis should be tested.

There are other instances where operant control of variability is adaptive. Some environments punish variability (e.g., many school classrooms), and most children can discriminate these from other situations (e.g., the game of hide-and-seek, where variability is reinforced). Environments that require brainstorming, problem solving, or creativity reinforce variability of behaviors. One attribute of successful art is the uniqueness of the artist's work. If an animal is to avoid predation or injury, it is functional for the animal to vary its behavior in the presence of specific predators or in specific environments (see Humphries & Driver, 1970; Serpell, 1982), and at least some aspects of this variability might be controlled by consequences. Operant variability is also functional in sports, gambling, war, or other competitive environments.

The interaction between elicited respondent variability and reinforced operant variability is an important area for future study. Respondent variability may set the boundaries within which reinforcing contingencies control operant variability. For example, both very low and very high densities of food to a hungry animal may prohibit high behavioral variability despite reinforcement of that variability (see Experiment 3). Alternatively, reinforcement of variability may extend the boundaries of elicited variation.

Reinforcement of variability raises theoretical problems. How can behavioral variation be reinforced if reinforcement increases the probability of those responses that produced it, thereby necessarily increasing response stereotypy (e.g., Schwartz, 1982a)? There are at least three possible ways in which reinforcement can serve both to increase the probability of prior responses and increase response variability. First, a schedule of reinforcement may rapidly condition and extinguish response patterns, thereby only apparently reinforcing variable sequences. It is the intermittent extinction, according to this interpretation, that elicits the variability, a respondent effect. However, when we attempted to train a single, fixed eight-response sequence in the stereotypy component of Experiment 6, we were unsuccessful. Despite hundreds of reinforcements over many sessions, the pigeons failed to learn. It therefore seems unlikely that they could learn an eight-response sequence after a single reinforcement. Furthermore, the present results did not show persistence of previously reinforced sequences.

A second hypothesis is that the pigeons learned a long sequence of responses or a rational strategy to meet the variability requirements. Experiment 4 showed that this, too, is unlikely.

The third interpretation, one supported by the present research, is that variability is a dimension of behavior much like other operant dimensions. Reinforcement does not necessarily lead to response stereotypy. Variability is as susceptible to control by reinforcement as are frequency, force, duration, location, or topography. But this does not imply that the existence of variability depends on its reinforcement. As indicated above, behavioral variability is a respondent consequence of environmental events as well as an operant progenitor. Furthermore, variable behavior must precede its consequences. The following analogy may be useful: The pigeon enters the operant conditioning experiment with a class of behaviors described as *pecking* already intact. When the experimenter shapes key pecking, the pecking response is not being trained. Rather, 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 enters our experiments. The rapidity with which inexperienced pigeons acquired variable performance under the initial Lag 50 condition in Experiment 5 supports this view. Turning on or off a variability generator may be under the control of reinforcement, but the variability generator is not itself created through reinforcement. An animal may be born with the variability generator intact.

The present results also raise a methodological issue. Skinner has argued for the ultimate predictability of behavior (Skinner, 1971) and therefore for the study of highly controlled behaviors (Skinner, 1984). In fact, operant conditioning studies have emphasized highly controlled acts (see Schwartz, Schuldenfrei, & Lacey, 1978), but a complete analysis of behavior must include analyses of the variability of behavior, variability maintained through respondent influences as well as that directly engendered and maintained by reinforcing consequences. It may be impossible to predict or control the next instance of a variable behavior, but lack of prediction and control should not rule out study. When behaviors are variable, experimental analyses can determine whether the variability is noise (i.e., experimental or observational error), under respondent control, or under operant control. Experiments can define the class or classes from which the individual instances of behavior are selected, the conditions under which the variability will occur or not, the variables controlling onset and offset of the variability, and the functions served by the variability. Operant analysis must not limit itself to predictable and controllable behaviors. Doing so ignores an essential characteristic of operant behavior.

Finally, we note one sociopolitical implication. *Freedom* (cf. Carpenter, 1974; Lamont, 1967) often means the absence of constraining contingencies (no gun to the head) and the presence of a noncontingent benign environment, one where adequate food, shelter, and so forth, are available independent of any particular behavior. However, operant contingencies may also be crucial. To maximize freedom, an animal or person must have a wide variety of experiences (Catania, 1980; Rachlin, 1980). If, for example, to obtain food, an animal has always entered the same 1 of 10 cubicles, each cubicle containing a different type of food, and therefore has never experienced any but the 1 food, it makes little sense to say that there is free choice among the 10 foods. The present results suggest that diversity of choice is controlled by reinforcers contingent on diversity. Absence of variability-maintaining contingencies, such as in the yoked condition of Experiment 5, increases stereotyped behaviors and therefore limits experiences. If these speculations are correct, a laissez-faire environment will not secure freedom of choice. Despite the absence of aversive constraints and the presence of positive respondent influences (e.g., good food, clothing, and company), behavior may still be highly constrained. Contingencies that explicitly reinforce behavioral variability are necessary to maximize freedom.

References

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

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

Bryant, D., & Church, R. M. (1974). The determinants of random choice. *Animal Learning & Behavior, 22,* 245–248.

Carpenter, F. (1974). *The Skinner primer: Behind freedom and dignity.* New York: Macmillan.

Catania, A. C. (1980). Freedom of choice: A behavioral analysis. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 97–145). New York: Academic Press.

Crossman, E. K., & Nichols, M. B. (1981). Response location as a function of reinforcement location and frequency. *Behavior Analysis Letters, 1,* 207–215.

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.

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

Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In B. Etzel, J. Le Blanc, & D. Baer (Eds.), *New developments in behavioral research: Theory, method and application.* Hillsdale, NJ: Erlbaum.

Humphries, D. A., & Driver, P. M. (1970). Protean defense by prey animals. *Oecologia, 5,* 285–302.

Killeen, P. R. (1978). Superstition: A matter of bias, not detectability. *Science, 199,* 88–90.

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

Lamont, C. (1967). *Freedom of choice affirmed.* New York: Horizon Press.

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

Miller, N. E. (1978). Biofeedback and visceral learning. *Annual Review of Psychology, 29,* 373–404.

Neuringer, A. (1970). Superstitious key pecking after three peck-produced reinforcements. *Journal of the Experimental Analysis of Behavior, 13,* 127–134.

Neuringer, A. (1984). Melioration and self-experimentation. *Journal of the Experimental Analysis of Behavior, 42,* 397–406.

Neuringer, A. (1985). *"Random" performance can be learned.* Manuscript submitted for publication.

Notterman, J. M. (1959). Force emission during bar pressing. *Journal of Experimental Psychology, 58,* 341–347.

Piscaretta, R. (1982). Some factors that influence the acquisition of complex stereotyped response sequences in pigeons. *Journal of the Experimental Analysis of Behavior, 37,* 359–369.

Popper, K. R. (1968). *The logic of scientific discovery.* New York: Harper & Row.

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.

Rachlin, H. (1980). *Behaviorism in everyday life.* Englewood Cliffs, NJ: Prentice-Hall.

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. (1981). Control of complex, sequential operants by systematic visual information in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 31–44.

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

Schwartz, B. (1982b). Reinforcement-induced stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General, 111,* 23–59.

Schwartz, B., Schuldenfrei, R., & Lacey, H. (1978). Operant psychology as factory psychology. *Behaviorism, 6,* 229–254.

Serpell, J. A. (1982). Factors influencing fighting and threat in the parrot genus Trichoglossus. *Animal Behavior, 30,* 1244–1251.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century.

Skinner, B. F. (1971). *Beyond freedom and dignity.* New York: Alfred A. Knopf.

Skinner, B. F. (1984, May). *The future of the experimental analysis of behavior.* Paper presented at the meeting of the Association for Behavior Analysis, Nashville, TN.

Received July 13, 1984
Revision received February 26, 1985 ■