

PROBABILITY RELATIONS WITHIN RESPONSE SEQUENCES UNDER RATIO REINFORCEMENT¹

FRANCIS MECHNER²

Columbia University

Skinner's original description of the effects of fixed-ratio reinforcement schedules in terms of response-rate measures (14) has served as the point of departure for all such subsequent investigations. The two salient aspects of fixed-ratio performance that have thus far received the most attention are the high response rates toward the end of the inter-reinforcement period and the characteristic pause after the delivery of a reinforcement (1, 3, 5, 8, 10, 13).

Fixed-ratio schedules also have some behavioral effects that are not easily described in terms of response-rate changes. For instance, the internal cohesion of response sequences maintained by ratio reinforcement is higher than it is for interval reinforcement. The extinction pattern after fixed-ratio reinforcement provides some of the evidence for this property. This pattern is characterized by maximal response rates which are maintained until they give way to an abrupt cessation of responding. If additional responses appear, they appear in bursts, rather than at the intermediate response rates that emerge during extinction after interval reinforcement (14). Further evidence for such cohesion can be seen in Dews' (4) and also in Herrnstein and Morse's (7) pharmacological data, which indicate that behavior maintained under ratio schedules is more resistant to disintegration by drug action than is behavior maintained under interval schedules.

The concept of "internal cohesion" of response runs (the term run is used in the sense of sequence) will be defined in terms of the probability that the run will terminate-- a definition which is not inconsistent with common usage. Thus, the cohesion of a run would be high when the probability of its termination is low. A systematic investigation of this property of response runs would, therefore, involve a description of runs in terms of their probabilistic structure, i. e., in terms of the probability that the run will terminate at any point. Once a technique is available, the effects of various parameters can be investigated.

¹ This experiment is part of a thesis submitted to Columbia University in 1957 as partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to express his gratitude to Professors W. N. Schoenfeld, F. S. Keller, and R. Berryman for their suggestions and encouragement.

² Now at Schering Corporation, Bloomfield, New Jersey.

One of the first studies to have direct bearing on this property of fixed-ratio runs is reported by Ferster and Skinner (5). In their procedure, a pigeon worked on a randomly mixed 50- and 250-response, fixed-ratio schedule. On those runs on which the pigeon was scheduled to receive reinforcement after 250 responses, it still paused after approximately every fiftieth response. The beginning of such a pause can be regarded as the point where a run ends.

Moskowitz used a somewhat more general procedure for investigating the length of fixed-ratio runs that are not interrupted by the presentation of reinforcement. This procedure, too, is described in the report by Ferster and Skinner (5). It involved the interspersing of extinction periods of various lengths among the 50-response ratio runs in such a way that the animal was never given any cues correlated with the beginning of an extinction period. After protracted training, the following pattern emerged. On those runs for which reinforcement was withheld, the animal stopped responding abruptly after having made some number of responses over the stipulated 50, and generally made few if any additional responses during the remainder of the extinction period. At the end of each extinction period, responding was restored by the presentation of an S^D .

Since a pause must be interpreted as the appearance of some behavior other than that being recorded, and since run-termination was defined in terms of such competing behavior, the points in the runs where the pauses occur are by definition points of low internal cohesion. Thus, a description of runs in terms of the probability of a pause appearing at any point (calculated on the basis of a large number of runs) would, according to this definition, constitute a description of the internal cohesion of the runs. However, because of the rather exploratory character of both the studies described above, the data were collected in a way that makes the extraction of the relevant information difficult. Therefore, a procedure was designed to implement a more systematic analysis of the cohesive properties of ratio runs. This procedure features a response on a second manipulandum whose purpose is to bring the competing behavior, which would otherwise be indicated by a pause in responding, under direct experimental control. This second response, by providing a discrete indication of run-termination, makes possible a description of the structure of response runs in terms of the probability of their termination at any point.

Under this procedure, rats were trained in two-lever Skinner boxes on a schedule where reinforcement was delivered either after (a) the completion of N consecutive responses on lever A, or else after (b) the completion of a minimum of N consecutive responses on lever B followed by an additional response on lever A. A random programmer determined which of these two conditions prevailed on any run. The probability, for any run, that reinforcement is delivered immediately upon the completion of N consecutive responses on lever A (condition (a)) will be referred to as P .

This procedure makes possible the measurement of the probability of termination of fixed-ratio runs before as well as beyond the point R_N , where they would normally be cut short by the presentation of reinforcement. To accomplish this, however, reinforcement must be withheld on some of the runs so that these can continue beyond R_N without interruption. This must be done often enough to provide sufficiently reliable estimates of the probabilities of termination. The problem now arises as to how such an interspersing of unreinforced runs will affect the probability function being recorded. Under straight fixed-ratio conditions (the case

of the schedule where $P = 1.00$), there are no unreinforced runs. Is it possible, then, for this probability function to be determined for pure fixed ratio? This dilemma was resolved in the following way. The frequency with which the unreinforced test runs were interjected was treated as a parameter in its own right. (This parameter is, of course, P , the probability that reinforcement is presented immediately after R_N . It was investigated for values of P ranging from 0.00 to 0.75.) The properties of the probability function for the case approaching fixed ratio could then be estimated by extrapolating along this parameter to the value where $P = 1.00$. Experimentally, the fixed-ratio situation could be approximated as closely as desired, by increasing the value of P . The closer P is taken to 1.00, however, the larger is the volume of data that would have to be collected to obtain a reliable function. If P were taken at 0.99, for example, only one out of every one hundred runs would contribute to the data being collected. Thus, the maximum proximity of P to 1.00 is limited by this practical consideration.

APPARATUS

The three boxes in which the animals worked were all of the same design. They were 7 inches wide, 8 inches long, and 7 1/2 inches high; and they had milk-glass ceilings, aluminum walls, and steel-rod floors. In each box, two T-shaped levers projected about 1/2 inch into the cage through two 3/4-inch holes in the front panel. These holes were 1 1/2 inches above the floor level, and were spaced 4 inches from center-to-center. The cross pieces of the levers were 2 1/2-inch-long- and 1/4-inch-wide-brass strips which were screwed on to the main levers. Depression of a bar (for which about 15 grams of force was required) actuated a microswitch which was positioned in such a way that the lever reached the lower limit of its excursion about 1/3 millimeter below the point where it actuated the microswitch. The limit-to-limit excursion of a bar was about 2 millimeters.

Water reinforcements (about 0.01 cubic centimeter per presentation) were delivered in cone-shaped cups drilled into the ends of 3/16-inch brass rods. The animals had access to the cups through 3/8-inch holes drilled in 1/2- by 1-inch shields that were horizontally positioned at the bottom centers of the front panels. The hinged dipper arms were pulled into water troughs underneath them by solenoids. The dipper arms were normally in the up position.

On the other side of each box was a food bin, the top of which was also level with the floor rods. This gave the animals continuous access to Purina Laboratory Chow during the experimental sessions. Each box was enclosed in an aluminum ice box which was ventilated by a wall air jet.

The "random" program mentioned in PROCEDURE was programmed on a 44-point stepping switch which was advanced every time a reinforcement was delivered. The distributions were recorded on batteries of impulse counters.

SUBJECTS

The subjects were male albino rats of Wistar stock, between 4 and 6 months old at the beginning of the experiment. Purina Laboratory Chow was continuously available to them, but water was supplied only as reinforcement during 2 1/2-hour work sessions. Animals worked daily, and always at the same time of day throughout the experiment.

PROCEDURE

Experiment A

Six animals were first trained on a 16-response, fixed-ratio schedule on lever A. Then the following conditions were put into effect. On 50% of the runs, reinforcement was presented, as before, after the sixteenth response on lever A. These runs were randomly interspersed among the other 50% of the runs by means of a random programmer, on which reinforcement was not presented until the animal had made at least 16 consecutive responses on lever A followed by an additional response on lever B. In case the animal switched to lever B before having completed the stipulated 16 responses on lever A, it had to repeat the entire sequence. The animals were maintained on this schedule for 14 days.

Each animal was then trained under each of four values (4, 8, 12, and 16 responses) of the parameter N according to sequences which had these features: (a) each value of the parameter appeared twice-- once in the first and once in the second half of the sequence-- for 9 consecutive days each time; (b) no value was ever preceded by the same value in both determinations; and (c) a different sequence was used for each animal. Data are reported for only the last 5 days of each 9-day period. A 4-day stabilization period was observed throughout for the sake of uniformity and expediency.

Experiment B

Five other animals were given the same preliminary training as the subjects of Experiment A, with two differences: (a) 8, rather than 16, responses on lever A were required as the minimum for reinforcement; and (b) 0.75, rather than 0.50, was used as the value for P. Each animal was then stabilized under each of four values of P (0.00, 0.25, 0.50, and 0.75) according to sequences that had the same features as those used in Experiment A.

RESULTS

The probability relations within the runs were examined in each of two ways:

1. The probability that the run will be terminated, as a function of the number of responses already made in that run. This relationship will henceforth be referred to as the $p(R_{T_n} \text{ given } R_n)$ function. (See GLOSSARY for the definitions of all symbols used.) The probabilities constituting this function were calculated by dividing the total number of times the animal switched to lever B immediately after R_n by the total number of times it reached R_n .
2. The absolute probability of a run of length n. This distribution will be referred to throughout as the $p(R_{T_n})$ distribution. Its probabilities were calculated by dividing the total number of times the animal switched immediately after R_n by the total number of runs.

The $p(R_{T_n} \text{ given } R_n)$ functions exhibited the following characteristics for all combinations of conditions investigated. The probability of R_T begins to rise several responses before N, rises most rapidly right after N, and then reaches a maximum. In all cases, the subsequent drop persisted as far out as the functions could be determined.

The $p(R_{T_n})$ distributions were always bell-shaped and roughly symmetrical, with their modes and medians various distances above N. Increases in the medians

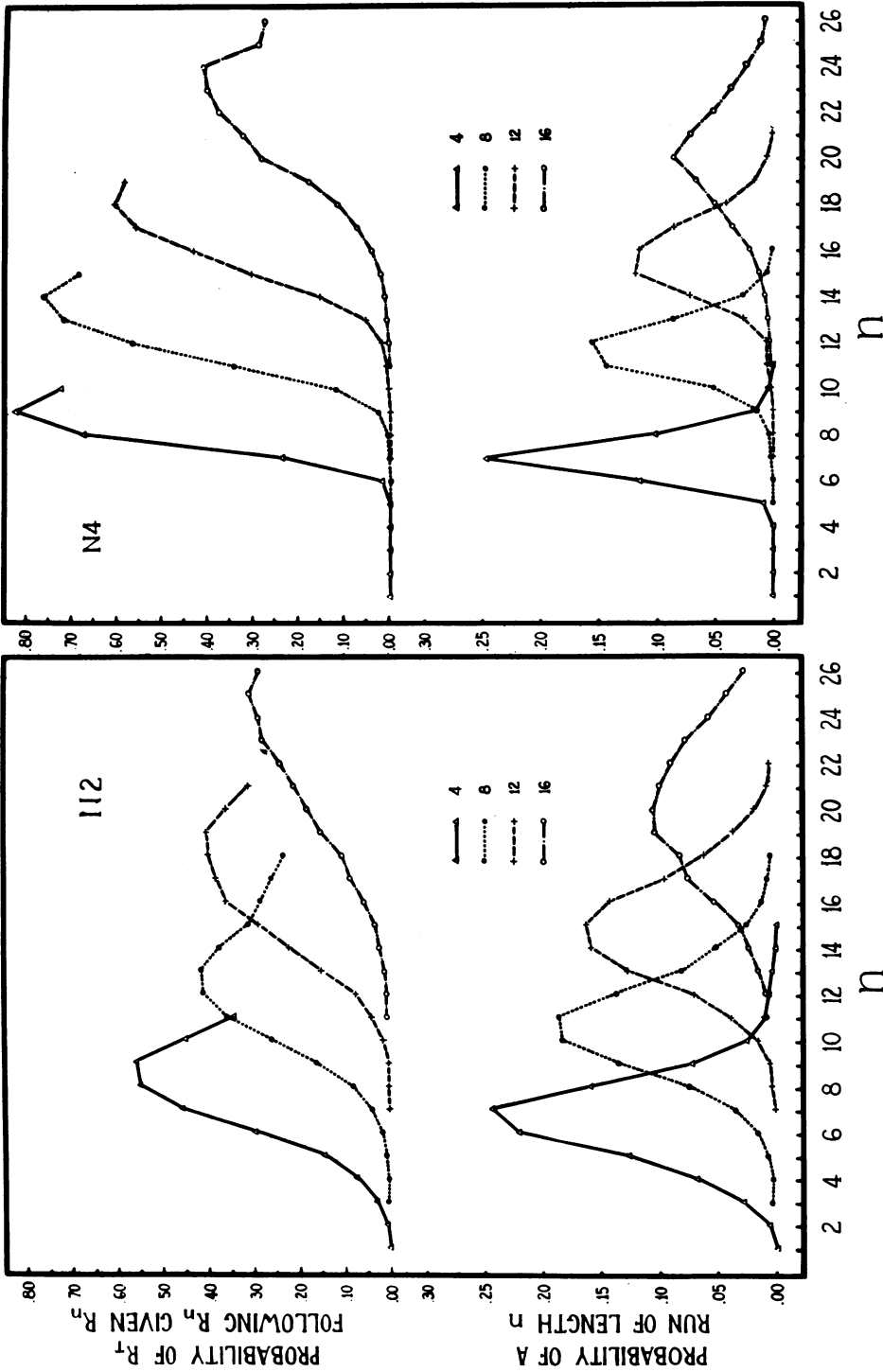


Fig. 1. Each of the two sets of coordinates shows the data of one animal. For each one, the four upper curves show the probability of switching to lever B as a function of the number of responses already made on lever A; and the lower four curves show the relative frequency of runs of various lengths. The parameter, whose values are indicated on the graphs as 4, 8, 12, and 16, is the number of responses required on lever A before the response on lever B will be reinforced. All data are based on those runs where the animal was required to switch.

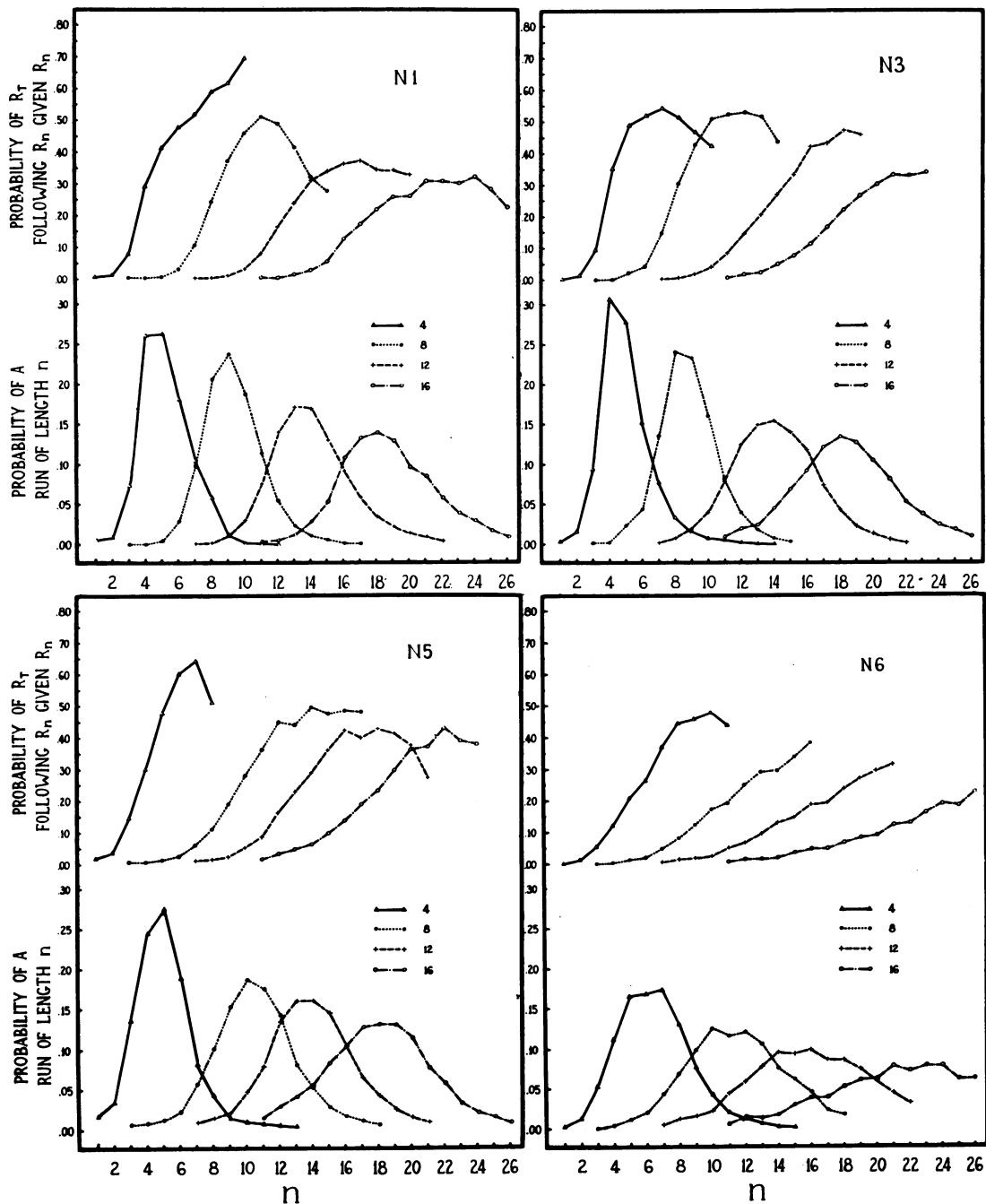


Fig. 2. Each of the four sets of coordinates shows the data of one animal. For each one, the four upper curves show the probability of switching to lever B as a function of the number of responses already made on lever A; and the lower four curves show the relative frequency of runs of various lengths. The parameter, whose values are indicated on the graphs as 4, 8, 12, and 16, is the number of responses required on lever A before the response on lever B will be reinforced. All data are based on those runs where the animal was required to switch to lever B for reinforcement.

due to the experimental variables were always accompanied by increases in the variances, although the mean was generally not equal to the variance (Fig. 1, Subject N4 especially); thus, the Poisson distribution probably would not provide a good fit.

Experiment A

The effect of increasing N on the $p(R_{T_N} \text{ given } R_N)$ functions, shown for each of six animals in the upper halves of the graphs in Fig. 1 and 2, is: (a) a displacement of the function along the abscissa in such a way that the steepest portion of the function always falls in the vicinity of N ; (b) a flattening of the steepest portion of the function; and (c) a lowering of the maximum value of $p(R_{T_N} \text{ given } R_N)$.

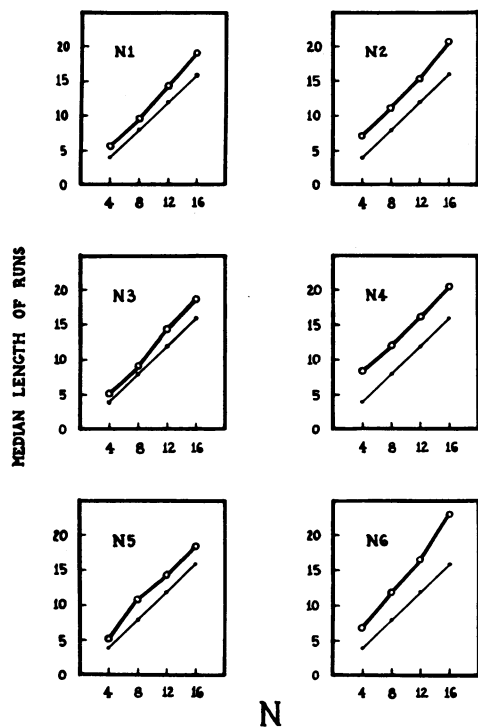


Fig. 3. Median lengths of runs as a function of N (the number of responses on lever A required before reinforcement can be obtained). These medians are based on the distributions shown in Fig. 1 and 2. The filled-in points indicate the values of N measured on the ordinate.

The effect of increasing N on the $p(R_{T_N})$ distribution was an increase in the medians as well as the variances. This effect can be seen in the lower halves of the graphs in Fig. 1 and 2. Figure 3 summarizes the effect by showing how the medians increase with increasing N .

Experiment B

The effect of increasing P is qualitatively similar to the effect of increasing N , and is shown in Fig. 4 and 5. The flattening of the $p(R_{T_N} \text{ given } R_N)$ function is, however, clearly visible only for Subjects P2 and P5. The decrease in the maximum value of $p(R_{T_N} \text{ given } R_N)$ is also not quite so striking as it is for the N parameter.

The performances of Subjects P1 and P3 were rather unstable. These animals used different techniques of executing their runs at different times. Sometimes, they gave consistently long, and at other times consistently short, runs under the same experimental condition. Subject P1 settled on the long-run type of performance after the third experimental condition, and maintained that performance for the balance of the experiment. Since the data from P1's two types

of performance could not be pooled, the results of the first and second determinations for $P = 0.25$ and 0.50 are presented separately in the upper half of the left-hand graph of Fig. 4. With Subject P3, however, not every type of performance was represented under each of the four values of P , so that the data were not presented.

The response rates within runs, which ranged from 3 to 10 responses per second, varied widely among the animals, but were quite consistent for any one animal,

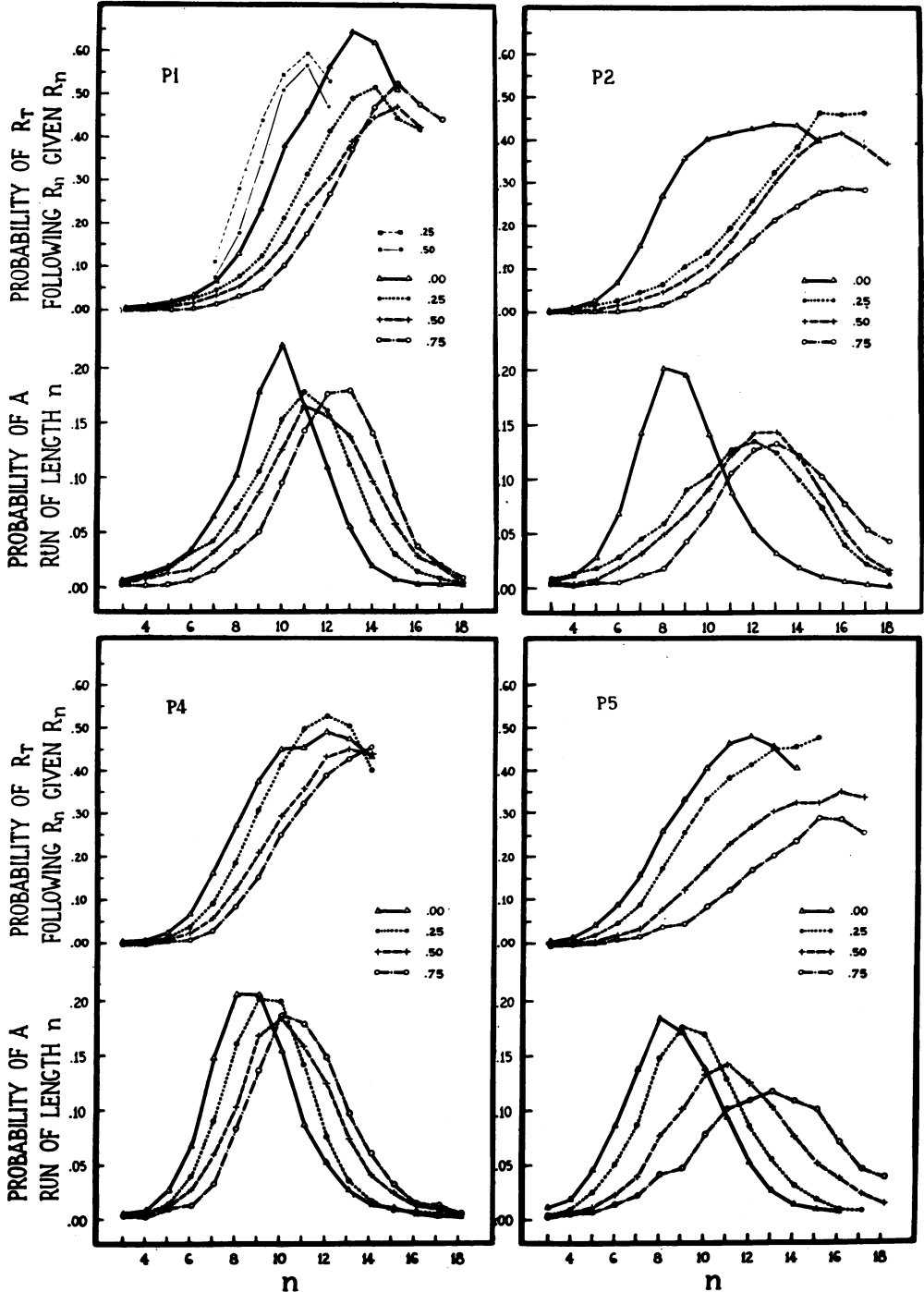


Fig. 4. Each of the four sets of coordinates shows the data of one animal. For each one, the four upper curves show the probability of switching to lever B as a function of the number of responses already made on lever A; and the lower four curves show the relative frequency of runs of various lengths. The parameter, whose values are indicated on the graphs as 0.00, 0.25, 0.50, and 0.75, is the probability that reinforcement is delivered immediately upon the completion of eight consecutive responses on lever A.

regardless of the value of N or P . The response rates on lever B, on the other hand, ranged from 4 to 14 per minute.

DISCUSSION

The function relating the probability of run-termination ($p(R_{Tn} \text{ given } R_n)$) to the point in the run reached (n) was offered as a description of the internal cohesion of the runs. Most of the following discussion will be devoted to the problem of explaining the effects of the parameters N and P upon the shape of this function.

These effects will be analyzed in the light of each of two not incompatible formulations, each one accounting for a somewhat different aspect of the results. The first formulation is built around the concept of an animal's discrimination of its own behavior, and owes its most explicit statement to Skinner (15), although it was originally suggested by Sheffield (11). The second one appeals to the notion of response topography variation, and was first proposed by Schoenfeld in 1950 (9).

An analysis of the behavior under the present schedule as a discrimination based on response-produced stimuli has to be predicated on the assumption that the response-produced stimuli are isomorphically coordinated with the responses. Once this assumption has been made, the responses on lever A can be considered the stimulus dimension for such a discrimination.

The first discrimination to be discussed is the one which is seen in its pure form in the $P = 0.00$ case. As the animal continues responding on lever A, the response-produced stimulus situation becomes increasingly favorable, that is, an ever-stronger S^D , for switching to lever B. The $P = 0.00$ functions shown in the upper halves of the four graphs in Fig. 4 illustrate the increasing probability of R_T with increasing n .

For values of P other than 0.00, the responses on lever A have a second source of strength: they are sometimes reinforced directly. This added contingency results in the superimposition of a second discrimination upon the one just discussed. Its effect is to give the responses prior to R_N added power as S^D 's for further responding. The amount of this increment would be greatest just before R_N (the point where reinforcement is actually delivered when it is forthcoming), and would diminish beyond that point. Thus, the degree of encroachment of the strength of the lever A response upon the strength of the lever B response should increase beyond

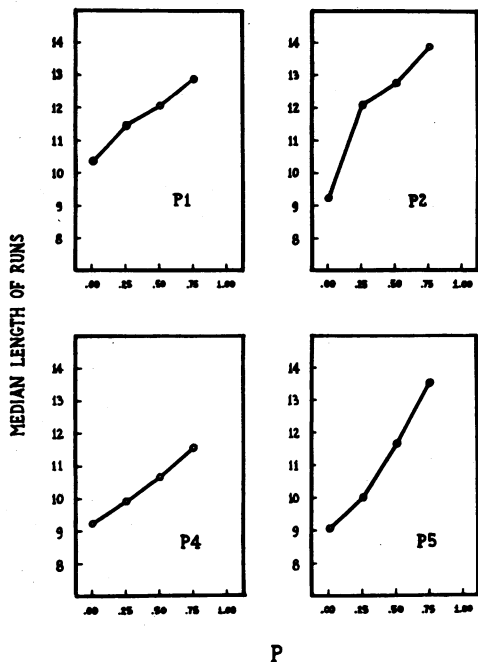


Fig. 5. Median lengths of runs as a function of P (the probability that reinforcement is delivered immediately upon the completion of eight consecutive responses on lever A). The medians are based on the distributions shown in Fig. 4.

N. In other words, when n is still close to N , the competition due to this added source of strength of the lever A response should syphon off a larger percentage of the probability of switching than when n is farther beyond N . This analysis is supported by the data of all four animals for which the effect of the parameter P is presented (upper halves of the graphs shown in Fig. 4). The effect of increasing P is, of course, to increase the magnitude of the effect due to this second source of strength of the first bar response: the higher the probability that it will be directly reinforced, the greater the extent to which it competes with the switching response. This second discrimination is never really isolated as a component of the probability functions, but is always superimposed upon the effect of the first discrimination. In order to isolate the behavior due to the second discrimination, P would have to be equal to 1.00 (straight fixed ratio), a condition that cannot be investigated, as these data can be collected on only those runs for which reinforcement is withheld. (The probability of this condition prevailing on any run, it will be recalled, is $1 - P$.) The shape of the function that would result if P could be taken at 1.00 can be estimated by extrapolation, as was discussed earlier. The complement of the resulting function would show the probability that the animal will continue responding on the first bar, and would, therefore, correspond to the pure case of the second discrimination.

The shape of the probability functions can also be explained by considering possible variations in response topography. During a run, the animal possibly could change its way of executing the individual responses on lever A gradually until it reaches that topographic variant which constitutes the proper S^D for switching to lever B. This S^D could be the tactual, kinaesthetic, and visual stimulus compound that results from the execution of that variant of the response. Although no record of response topography was obtained in these experiments, this account is supported by informal observations. One rat pressed lever A with its right paw while describing a semicircle on the wall with its left paw. Upon reaching the end of this semicircle, it would switch. This pattern was repeated run after run. Another rat pressed lever A with its teeth. During the run, it would gradually lower its head until its teeth slipped off the lever, at which point it would switch.

These observations also explain certain aspects of the present data that would otherwise be difficult to handle. One is the variation from one determination to the next (there were two determinations for each experimental condition), and also from one animal to the next, in the relative positions of the probability functions along the abscissa. An instance of the determination-to-determination variation can be seen in the $N = 4$ functions for Subject N1. The slight inflection in the $N = 4$ function shown in the upper half of its graph is the result of pooling the first and second determinations. Another instance appears for Subjects P1 and P3 (Fig. 4), where the discrepancy between the results of the first and second determinations was so large that the data from the two determinations could not be pooled, as was mentioned in RESULTS. These variations in the positions of the functions are at first surprising in view of the obviously very high reliability of these functions. The displacements could, however, result in the following manner. When an animal is shifted from one value of N to another (say, a higher one), the particular sequence of topographic variants of the response that constituted its run on lever A is no longer effective and undergoes extinction. During this extinction, some longer runs will emerge, one of which will eventually be of sufficient length to meet the new requirement. Thus, a new chain evolves. Its final attributes will depend upon the fortuitous characteristics of the runs that were its ancestors. For example, the

average length of the new chain will depend upon the amount by which the lengths of the runs from which it is descended were in excess of the new requirement. The new chain on which the animal stabilizes will, therefore, be highly reproducible from day to day so long as the animal is maintained under the same set of conditions, but will not necessarily be recoverable after the animal has been shifted to another set.

The maxima in the $p(R_{Tn} \text{ given } R_n)$ functions can be accounted for in two ways, neither of which possesses much theoretical significance. The first possibility is that the maxima are a consequence of small fluctuations in the steepness of the functions. Since the probabilities are conditional, i. e., are computed on the basis of the number of times that point in the run was actually reached, the farther out along n the probability is taken, the more heavily will the flatter functions be represented; the flatter the function, the larger the number of long runs on which it is based. The maxima are, therefore, the points where the steepness of the short-run functions begins to be offset by the heavier representation of the long-run functions.

The second possibility is that on rare occasions, the animal, for one reason or another, does not carry a run to completion (as when it does not depress lever B sufficiently to actuate the microswitch, or is interrupted in the middle of a chain by slipping on the floor rods) and starts the entire sequence from the beginning. Such occurrences would, of course, be recorded as unusually long runs, and would, despite their rarity, affect the shape of the function in the high- n region where they have relatively little competition from authentic runs.

SUMMARY

A procedure was developed for the purpose of investigating the internal cohesion of response sequences maintained on fixed-ratio reinforcement. Under this procedure, rats were trained in a two-lever Skinner box on a schedule wherein water reinforcement was delivered either upon the completion of N consecutive responses on lever A, or else upon the completion of a minimum of N consecutive responses on lever A followed by an additional response on lever B. A random programmer determined which of these two conditions prevailed on any run.

The two main parameters investigated were N (the minimum number of responses required for reinforcement) and P (the probability that the animal is reinforced immediately upon the completion of the N responses on lever A). The values of N used were 4, 8, 12, and 16, and the values of P were 0.00, 0.25, 0.50, and 0.75.

The following two functions were calculated for the data of each animal for each of the four values of the variable to which it was exposed: (a) the probability of switching to lever B as a function of the number of responses already made on lever A, and (b) the frequency distributions of lengths of runs. The effect of increasing N was a shift in both of these functions upward along the abscissa, in such a way that the sharpest rise in function (a) and the median of distribution (b) always fell slightly above the stipulated value of N . The effect of increasing P was qualitatively similar to the effect of increasing N .

GLOSSARY OF TERMS

Run	A succession of responses on lever A ending with a response on lever B.
-----	---

Run-length	The number of consecutive responses made on lever A before a response is made on lever B. Runs of length zero are excluded from the data.
N	The minimum number of responses on lever A required for reinforcement.
P	The probability that reinforcement will be delivered immediately upon the completion of N consecutive responses on lever A.
n	The number of responses in a run.
R _N	The Nth response in a run.
R _n	The nth response in a run.
R _T	The response on lever B which marks the termination of the run.
R _{Tn}	A response on lever B (i. e., an R _T) immediately after R _n .

REFERENCES

1. Anger, D. The dependence of inter-response times upon the relative reinforcement of different inter-response times. *J. exp. Psychol.*, 1956, 52, 145-161.
2. Antonitis, J. J. Variability of response in the white rat during conditioning and succeeding extinction and reconditioning. Unpublished doctoral dissertation, Columbia Univer., 1950.
3. Boren, J. J. Response rate and resistance to extinction as functions of the fixed ratio. Unpublished doctoral dissertation, Columbia Univer., 1955.
4. Dews, P. B. Studies on Behavior. I. Differential sensitivity to pentobarbital of pecking performance in pigeons depending on the schedule of reward. *J. Pharmacol. & Exper. Therap.*, 1955, 113, 393-401.
5. Ferster, C.B., and Skinner, B.F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.
6. Graham, C.H. Behavior, perception, and the psychophysical methods. *Psychol. Rev.*, 1950, 57, 108-120.
7. Herrnstein, R.J., and Morse, W.H. Selective action of pentobarbital on component behaviors of a reinforcement schedule. *Science*, 1956, 124, 367.
8. Lagmay, A. Pacing technique in the study of some schedules of reinforcement. Unpublished doctoral dissertation, Harvard Univer., 1955.
9. Schoenfeld, W.N. On the difference in resistance to extinction following regular and periodic reinforcement. C. E. A. B. Note No. 20, February 27, 1950 (mimeographed, issued from Indiana University, Department of Psychology, Bloomington, Indiana).
10. Schoenfeld, W.N., Cumming, W.E., and Hearst, E. On the classification of reinforcement schedules. *Proc. Nat. Acad. Sci.*, 1956, 42, 563-570.
11. Sheffield, V.F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, 39, 511-575.
12. Sidman, M. Time discrimination and behavioral interaction in a free operant situation. *J. comp. physiol. Psychol.*, 1956, 49, 469-473.

13. Sidman, M., and Stebbins, W. C. Satiation effects under fixed-ratio schedules of reinforcement. *J. comp. physiol. Psychol.*, 1954, 47, 114-116.
14. Skinner, B. F. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
15. Skinner, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 57, 193-217.