Psychonomic Science, 1, 1964, pp. 361-362.

Probability learning of rats in continuous-time experiments[1]

Eileen B. Karsh, *DREXEL INSTITUTE OF TECHNOLOGY*
Patrick Suppes, *STANFORD UNIVERSITY*[2]

Abstract

Probability matching behavior in time was shown by rats under two non-contingent reinforcement schedules (.50 and .625) with correction plus shock procedure but not with non-correction. Sequential data indicated that the maximum effect of reinforcement occurred about 6 sec. later and that only two preceding reinforcements were effective in determining response probability.

Problem

During the past ten years there have been many studies on probability learning in humans, other mammals, and animals lower on the phylogenetic scale (Calfee, 1963; Estes, 1964). The prime interest in much of this research has been whether, for the given species and experimental conditions, probability matching occurs. In the case of the rat, it has proved difficult to find conditions where clear evidence of probability matching is obtained for individual subjects during long-term experiments (Calfee, 1963; Uhl, 1963).

The present paper presents results for probability learning with rats in an experimental setting not previously used for this purpose, a symmetric shuttlebox. In continuous-time experiments we consider the response-state of the organism at all times, rather than discretely emitted responses. Thus, the subject is always in a response-state, for it must be on either the left or the right side of the shuttlebox. In order to determine whether probability matching could be obtained, we varied experimental conditions and examined mean asymptotic response probabilities for individual subjects under each condition.

In addition, the evaluation of sequential statistics was expected to provide basic information about reinforcement. Would the effects of reinforcement be clearly observable? Specifically, when does the maximum effect of a reinforcement occur in the interval following its delivery? Are the effects of reinforcement accountable simply in terms of the last reinforcement, or are there long-term effects reaching back through preceding reinforcements? Finally, if sequential dependencies are shown, are negative recency effects also obtained, as in studies of human probability learning?

In the present paper we are primarily interested in presenting experimental results; thus no effort has been made to provide a quantitive analysis of the data from the standpoint of a mathematical theory of learning, although such analysis is now in progress and will be reported later.

Method

Subjects were six adult male rats on food deprivation—at 80% normal weight. The apparatus was a standard shuttlebox (30 in long x 5 in wide x 6 1/4 in high) with a grid floor and a low barrier (1 1/4 in) in the center. Food cups were mounted on both ends of the box. The shuttlebox was dimly illuminated from above and was enclosed in a sound resistant outer box, which contained a viewing panel for observation. Shock was delivered from a matched-impedance source with 150 K series resistor.

Rats were first magazine-trained to take 45 mg food pellets from both feeders. There was no difficulty in getting animals to move freely across the box.

Throughout the experiment, a fixed interval outcome presentation was used. Every 20 sec. a light came on briefly (1/2 sec.) from below the grid, on one side of the box. Two hundred such presentations were given during a daily session, which lasted slightly more than 1 hr. These events were programmed in random series on a tape reader. The location of the rat was recorded throughout the session, and a summary measure, the proportion of time spent on the left (A_1) side, was recorded for each session.

Initially a non-correction procedure (NC) was used. If the rat was on the side where the light came on, a pellet was delivered; if the rat was not on the side where the light appeared, no reinforcement was given. Later the procedure was changed to correction plus shock (C + S). Under this procedure, if S was on the wrong side when the light came on, it received a fairly weak shock (2 sec. at 50 volts). By crossing over to the other side before the next light S could get a pellet and, by crossing rapidly, could also escape from the shock.

A number of different conditions were used consecutively for all Ss, so that each rat was run daily over a period of about 7 months. With the non-correction procedure the proportion of E_1 events (light on the left side), π, was initially .625 for two rats and .375 for two rats. Next, the more frequent side was reversed (π = .375 or .625), and then the schedule was changed to π = .50. Two additional rats were started under this condition (π = .50 NC). Then the correction plus shock procedure was adopted, with π = .50 and shock at 50 volts. Shock was then increased to 55 volts but later changed back to 50 volts, with π = .50. Finally, the schedule was changed to π = .625 (or .375), with shock at 50 volts.

Mean Asymptotic Results

After 15-20 days under the first condition (π = .625 or .375 NC), Ss showed absorption, i. e., spent almost all their time during each session on the more frequently reinforced side. When the schedule was reversed to π = .375 or .625 NC, Ss absorbed on the other side after 10-15 days. When the schedule was changed to π = .50 NC (shown on the left in Fig. 1), two of the four rats (No. 151 and No. 156) continued to show decided position preferences although it was no longer gainful. The other two rats of the original group (No. 150 and No. 164) and the two later ones (No. 170 and No. 171) did not deviate much from .50, considering performance for all 25-35 days. However, there was relatively large day-to-day variability. With the institution of the more powerful correction plus shock

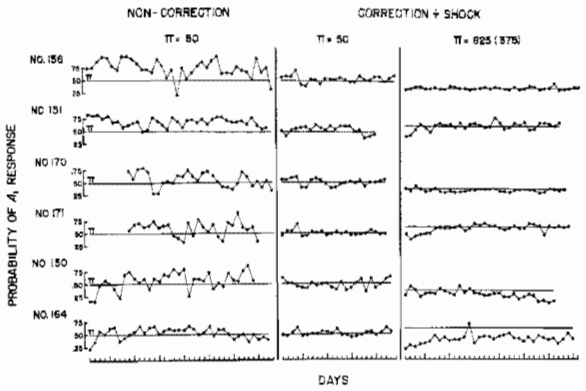

Fig. 1. Mean proportion of time in A_1 response-state during daily sessions for individual Ss under 3 conditions: π = .50 NC, π = .50 C + S, and π = .625 (.375) C + S.

procedure, all six rats showed probability-matching for the 20-25 days (i. e. 4000-5000 outcome events), under π = .50 with shock at 50 volts. A decrease in variability was also evident, as shown in the center of Fig. 1. When shock was increased to 55 volts, two rats (No. 150 and No. 164) began to show position preferences which persisted throughout the remainder of the experiment. The other four rats continued to probability-match throughout the C + S conditions with π = .50 and also with π = .625 (.375), as shown on the right in Fig. 1. Thus all six rats showed probability-matching under the .50 schedule and four showed good probability matching under the .625 schedule for a total of about 100 consecutive days or 20,000 outcome events.

Sequential Statistics

Because outcomes occurred at fixed intervals of 20 sec., it was relatively easy to examine sequential statistics. Data were analyzed for all six rats for 15 days under the .50 schedule (50 volts).

Figure 2 shows the probability of an A_1 (left) response, given that an E_1 outcome (light on left) occurred; the figure also shows a curve similarly conditionalized on an E_2 outcome. Each curve gives the mean probability that S was in an A_1 response-state as a function of seconds following the last outcome (E_1 or E_2).

It is clear that the effects of reinforcement are very strong. The complete separation of the two response curves, following different outcomes, is about as great as one could expect.

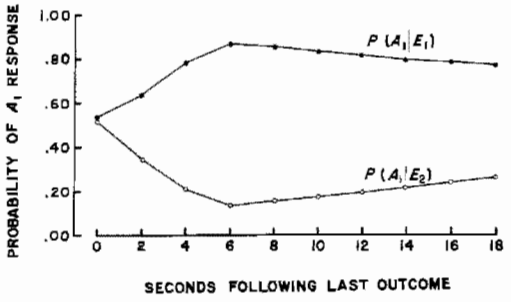

Fig. 2. Mean probability of an A_1 response-state after an E_1 or an E_2 outcome, as a function of seconds following the last outcome.

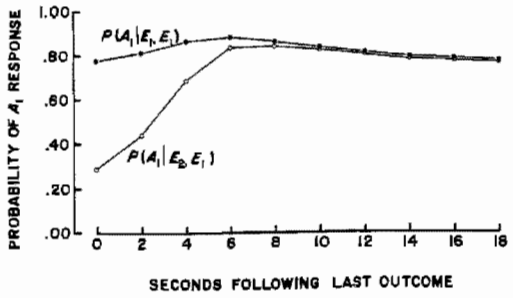

Fig. 3. Mean probability of an A_1 response-state after two E_1 outcomes (E_1, E_1) or an E_1 preceded by an E_2 (E_2, E_1) as a function of seconds following the last outcome.

It is interesting to note that the maximum effects of reinforcement do not occur until 6 sec. after the outcome is presented. Naturally following from such an observation is the question of a possible variation in this temporal lag when additional sequential statistics are considered. In Fig. 3 the conditional response curves for two preceding outcomes are shown. In the top curve the maximum effect of reinforcement occurs after 6 sec. and in the second one after 8 sec. The results shown here are correlated with those shown in Fig. 2 regarding the time of maximum effect, but the essential point is that the consideration of more conditional statistics does not lead to much spread in the distribution of maxima. In fact, if we look as far back as four preceding outcomes, the range of occurrence of the maximum is between 4 and 8 sec. (e.g., it is 4 sec. for the sequence E_1, E_2, E_2 E_2 and 8 sec. for E_2, E_2, E_1, E_2). The rather long temporal lags following the reinforcing outcome are surprising. As far as we know, this kind of measurement of the time it takes for reinforcement to show its maximum effect has not been reported proviously.

Figure 3 also shows evidence of sequential dependencies, because of the clear separation of the response curves conditionalized on E_1, E_1 outcomes and E_2, E_1 outcomes. The last outcome was the same (E_1) in both cases, but the effect of the preceding outcome made a considerable difference. This effect was not observed for three preceding outcomes; in fact, the effect of the third preceding outcome was very slight indeed. Finaliy, considering the 16 possible combinations of four preceding outcomes, very little effect of the earliest outcome was noticeable. In particular, there was no tendency at all for a negative recency effect to appear.

In conclusion, it may be remarked to those familar with stimulus sampling models that the comparison between the E_1, E_1 and E_2, E_1 curves suggests that the estimate of the conditioning parameter c will be somewhere in the neighborhood of .8. The differences in the maxima of these two curves and the fact that both maxima are less than 1 show, of course, that c will not be precisely 1. More detailed questions about the estimation of parameters and the fit of models will be delayed for a second paper.

References

CALFEE, R. C. Long-term behavior of rats under probabilistic reinforcement schedules. Tech. Rep. No. 59, Institute for Mathematical Studies in the Social Sciences, Stanford Univer., 1963.

ESTES, W. K. Probability Learning. In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

UHL, C. N. Two-choice probability learning in the rat as a function of incentive, probability of reinforcement, and training procedure. *J. exp. Psychol.*, 1963, 66, 443-449.

Notes

1. This research was supported by contract AF 49 (638) - 1253 and grant AF AFOSR 62384 between the Air Force Office of Scientific Research and Stanford University and by grant MH 08419 from the national Institute of Mental Health to Drexel.

2. We wish to thank Dr. Saul Sternberg for use of the apparatus. We also thank J. Seiler for circuit design and A. Kirschenstein, R. Franks and M. Deutsch for experimental assistance.