T-maze learning as a joint function of probability and magnitude of reward¹

Keith N. Clayton and Sally T. Koplin
VANDERBILT UNIVERSITY

Abstract

Rats were trained on a position discrimination. The experimental design involved simultaneously manipulating per cent and magnitude reward given for the correct response. Learning-rate was found to be a negatively accelerated growth function of reward magnitude for both of the reward schedules used. The results are interpreted as support for the treatment of reward magnitude as a special case of reinforcement probability.

Problem

In a previous study (Clayton, in press) three groups of rats received either 1, 2, or 4 pellets for the correct position response in a T-maze and no reward for an error. Analysis of the resulting performance suggested the following relationship between learning-rate and reward magnitude:

$$(1) J_r = 1/M [1 - (1 - MJ_1)^r] \quad (MJ_1 \leq 1),$$

where J_r is the learning-rate parameter associated with r reward pellets. Equation (1) says, generally, that J_r is a negatively accelerated growth function of J_1 , where M is a parameter effecting the growth-rate and asymptote of the curve.

The present study was stimulated by an assumption regarding reward magnitude made recently by Estes & Suppes (1959). Following a Guthrie treatment of reward magnitude and distinguishing between reward and reinforcement, Estes & Suppes have suggested reward magnitude is but a special case of reinforcement probability in that an increase in magnitude merely increases the probability of the reinforcing event. Within this framework Equation (1) may be conceived as the relationship between learning-rate and reinforcement probability. However, a second variable affecting reinforcement probability should be reward probability, or reward schedule. In the previous study the correct response was always rewarded. The present study sought to determine the effect on the relationship between learning-rate and reward magnitude when the correct response was sometimes nonrewarded. The hypothesis tested was that if Equation (1) relates learning-rate to reinforcement probability, it should hold (with possible changes in M) regardless of whether the correct response was always rewarded.

A second purpose of this study was to provide an independent check on Equation (1) with different reward magnitudes to make certain the obtained function is not specific to the particular amounts used previously. For these reasons reward magnitude was factorially manipu-

lated with reward probability, resulting in a 3×2 design in which the magnitudes were 1, 3, and 6 pellets and the per cents were 100 and 50.

Method

The Ss were 96 female albino rats, purchased from Holtzman Rat Co., 91 days of age at the beginning of acquisition. The apparatus was the single-unit T-maze described elsewhere (Clayton, in press). The procedural details followed closely those utilized in the previous study. The Ss were given nine days of prehandling, which consisted of adjustment to handling, to the daily feeding schedule, and to the reward pellets. No preliminary exposure to the apparatus was given. Four training trials were given daily except the first two days during which two trials were given daily. All trials were free-choice trials. After S's choice, a wooden door prevented retracing and 15 sec. confinement in the goal box was given regardless of choice. Reward, when given, was either 1, 3, or 6 45 mg Noyes food tablets as required by S's treatment. No reward was ever given for an error and correction was never permitted. For Ss in the 50% groups, four reward schedules were utilized and a fourth of each 50% subgroup was randomly assigned to each reward schedule. Each 50% schedule was strictly random, i. e., no restrictions were placed on the reward events, which were determined from a table of random numbers.

A left turn was designated correct for half the Ss in each of the subgroups, the right side was positive for the remainder. In addition, the rats were run by two experimenters and this variable was similarly factorial to magnitude, per cent, and position, so that the final design was a $3 \times 2 \times 2 \times 2$ factorial with four rats per cell.

All rats received at least 31 days training and rats which had not reached a criterion of learning by 31 days were continued until the criterion was reached. The criterion was three consecutive errorless days (12 trials).

Finally, the intertrial interval was approximately 15 sec. and the daily food ration was given 30-45 min. after the last trial of the day.

Results and Discussion

Table 1 presents the mean errors-to-criterion data for the six major groups of the study. Analysis of variance of these data revealed significant F's for magnitude ($F=40.7$; $2/72$ df; $p < .001$) and per cent ($F=37.4$; $1/72$ df; $p < .001$) and no other main effects or interactions were significant.

Table 1. Mean Errors to Criterion ($\bar{T}$) by the Six Experimental Groups

Magnitude		Per Cent	
		100	50
6	Mean	4.69	13.69
	Standard error of mean	.64	1.66
3	Mean	7.50	17.88
	Standard error of mean	.87	2.12
1	Mean (observed)	21.12	36.25
	Mean (predicted)	19.16	38.03
	Standard error of mean	3.26	3.33

To determine whether Equation 1 holds for the present data the following procedure, equivalent to that previously used, was taken. First learning-rates for the 6 and 4 pellet groups were estimated utilizing:

$$(2) \hat{J} = q_1 \frac{1}{\bar{T}},$$

where q_1 is the probability of making an error on trial 1, assumed a priori to be $1/2$, and $\bar{T}$ is the mean errors to criterion.

From Equation (1) it can be shown that

$$(3) M = \frac{2J_{3,0} - J_{6,0}}{(J_{3,0})^2};$$

thus M was estimated using the 6 and 3 pellet subgroups of each of the 100% and 50% groups. For the 50% condition the estimate of M was 24.81; for the 100% condition M was 6.00. Finally, the learning-rate parameter for the 6 pellet groups and the obtained M 's were substituted into Equation (1) to obtain the predicted learning rate for the 1 pellet-50% and 1 pellet-100% groups separately. These predictions, transformed into mean errors-to-criterion, are presented in Table 1. As can be seen the predictions fall within a standard error of the obtained mean.

It is concluded that the results are in excellent agreement with the hypothesis that the previously obtained function relating learning-rate to reward magnitude is substantially correct and that it applies to learning-rate, with appropriate parameter changes, regardless of the per cent reward schedule for the correct response. And this general finding is considered to be indirect support for Estes & Suppes' treatment of reward magnitude as a special case of reward probability. It is recognized this conclusion implies acceptance of the null hypothesis (see Binder, 1963; Grant, 1962)—but it should be noted that some care was taken initially to have a reasonably large number of experimental subjects.

Fig. 1. The learning-rate parameter, J , as a joint function of probability and magnitude of reward as suggested by the results of the present study. The circles were obtained from the data of the 6 experimental groups reported here.

Finally, the general picture beginning to emerge of the joint effect of reward magnitude and reward per cent on learning-rate is depicted in Fig. 1. This surface must be regarded as tentative at the present time and the degree to which it depends on the experimental procedures so far used is not fully understood. However, it utilizes the strong suggestion, resulting from the two estimates of M , that $1/M$ [the asymptote of Equation (1)] is a parabolic function of per cent—a result reminiscent of Restle's (1957) suggestion that cue validity is a parabolic function of reward probability.

References

- BINDER, A. Further considerations on testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychol. Rev.*, 1963, 20, 107-115.
- CLAYTON, K. N. T-maze choice learning as a joint function of the reward magnitude for the alternatives. *J. comp. physiol. Psychol.*, in press.
- ESTES, W. K., & SUPPES, P. Foundations of linear models. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 137-179.
- GRANT, D. A. Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychol. Rev.*, 1962, 69, 54-61.
- RESTLE, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, 64, 182-191.

Note

1. This study was supported by National Institute of Mental Health Grant MH-05576. The authors are indebted to William B. Harrison, Jr., who ran half the Ss. This paper was presented in part at the 1964 meetings of the Midwestern Psychological Association.