Acquisition of discrimination between rewards¹

MARGARET B. CAKMAK AND NORMAN E. SPEAR

RUTGERS UNIVERSITY

Three experiments tested the generality of Clayton's (1964) model for describing and predicting the rate of acquiring discrimination between rewards. With massed training, (Experiments I and II), the data could be adequately described by the model. However, when trials were spaced (Experiment III), trapping occurred; and results were inconsistent with the assumptions underlying the model.

When rats are trained in a T maze with *s* pellets for reward for a "correct" choice and *t* pellets for an "incorrect" choice (*s* > *t*), a learning rate parameter *J* may be predicted from the equation $J_{s,t} = J_{s,o} - J_{t,o}$, stating that the learning rate of *Ss* choosing between a larger and a smaller reward is equal to the difference between groups choosing between the larger reward paired with none and the smaller reward paired with none (Clayton, 1964). The first experiment below attempted to extend this model to a somewhat different situation.

EXPERIMENT I

This experiment investigated the acquisition behavior of three groups, comparable to groups employed by Clayton (1964): Group 4-2 (four pellets reward for the correct alternative and two for the incorrect), Group 4-0, and Group 2-0. Deviating from Clayton's procedures, the *Ss* were assigned their non-preferred side as correct, were rewarded with sucrose pellets rather than regular Noyes pellets, were maintained on a smaller daily food ration, and (on nonrewarded trials) found an empty food cup in the goal box. In all other major particulars, Clayton's procedure was followed as closely as possible.

Method

Thirty male and female Sprague-Dawley rats, aged 60-90 days, were trained on a position-discrimination task in the T maze described by Spear & Hill (1965, Experiment II). *Ss* were prehandled and given 52 free massed (15 sec. intertrial interval) acquisition trials at the rate of four per day.

Results

The probability of a correct choice for each group for each two-day block of trials (except Day 13) may be seen in the left portion of Fig. 1. As there were no differences in the results for males and females, the data were pooled for analysis. According to Clayton's model, a learning rate parameter *J* may be estimated from error data such that $\hat{J} = q_1 \frac{1}{\bar{T}}$, where $\bar{T}$ equals the group mean total errors in acquisition (52 trials in Clayton's and the present study), and *q*₁ represents the probability of an error on Trial 1. In Clayton's study, correct side was assigned randomly and *q*₁

Fig. 1. Probability of choosing the "correct" alternative in acquisition: Experiments I, II, and III.

was taken to equal .5. In the present study, *Ss* were run to their non-preferred side and thus *q*₁ was 1.00. Note, however, that the predicted relationship between $\hat{J}_{4,2}$ and the difference $\hat{J}_{4,0} - \hat{J}_{2,0}$ is independent of the value of *q*₁. That is, if $\frac{1}{\bar{T}_1} = \frac{1}{\bar{T}_2} - \frac{1}{\bar{T}_3}$, then the fractions $\frac{1}{\bar{T}_1}$ may be multiplied by any constant *q*₁ without altering the relationship.

Because not all *Ss* had attained 100 percent preference for the "correct" side and because the assumption of asymptotic choice probability of 1.00 is critical to the underlying single-operator linear model, least-squares curves of the form $y = m \log x + b$, were fit to the obtained error data from each group and extrapolated to the point where error probability was less than .012. These functions were then employed to estimate total errors to a strict criterion. These were 16.6, 25.7, and 36.4 for Groups 4-0, 2-0, and 4-2, respectively, and $\hat{J}_{4,0} - \hat{J}_{2,0} = .060 - .039 = .021 \approx .027 = \hat{J}_{4,2}$. A curve based on $\hat{J} = .021$ was generated and compared to the obtained error data of Group 4-2. The obtained scores did not deviate significantly from these generated data, $X^2(12) = 7.16$, $p > .50$, indicating a reasonably good fit with the predictions of Clayton's model.

EXPERIMENT II

Method

Experiment II involved a closer replication of Clayton's procedures and a much longer acquisition series in order to establish stable discrimination in all

groups. Forty-eight female Sprague-Dawley rats weighing 180-200 gm were prehandled and given 34 days of acquisition in the T maze described by Spear & Spitzner (1966). The remaining procedure was identical to Experiment I except that reward consisted of regular rather than sucrose pellets, the food cup was absent on nonrewarded trials, correct side was assigned randomly, and daily ration was 12 gm. In each of these changes, a more nearly exact replication of Clayton's procedure was effected. A randomly selected half of each group was given additional acquisition trials on Days 35-39.

Results

The probability of making a correct choice has been plotted for three-day blocks of trials in the middle portion of Fig. 1. Even after 39 days (156) trials, Group 4-2 was not exhibiting perfect performance. Thus, a least-squares line was fit to the choice probabilities for Group 4-2 from Day 13 through Day 39 and extrapolated to Day 60 when, had the linear trend continued, error probability would have reached .01. Using this extrapolation to estimate $\bar{T}_{4,2}$ and the obtained $\bar{T}_{4,0}$ and $\bar{T}_{2,0}$, Clayton's model yields a very close prediction; $\hat{J}_{4,0} - \hat{J}_{2,0} = .040 - .029 = .011 \approx .012 = \hat{J}_{4,2}$.

EXPERIMENT III

Method

Forty-eight naive female Sprague-Dawley rats weighing 180-200 gm were trained in the same maze as in Experiment II. The procedure was identical to that of Experiment I except that trials were spaced at least 5 min. apart in order to reduce spontaneous alternation and thus increase the rate of discrimination.

Results

The probability of a correct choice is plotted across two-day blocks of trials in the right-hand portion of Fig. 1. Groups 4-0 and 2-0 both readily achieved near perfect performance. Although Group 4-2 remained close to chance level, the curve is not really representative of that group's performance. In fact, four of sixteen 4-2 Ss always chose the incorrect side, and two others were performing below chance level at the end of acquisition. Some of the 4-2 Ss were also "trapped" on the correct side. That is, three of the 4-2 Ss made only one error in 52 trials and that was on the first trial (necessarily an error since choice on Trial 1 was used to define the incorrect side). No curve fitting was attempted since there was no reason to suppose that the trapped Ss would alter their preferences permitting the function to reach 100 percent choice of the larger reward.

DISCUSSION

In Experiments I and II, performance of Group 4-2 did not differ significantly from that predicted by Clayton's equation. It should be noted that in the present experiments, learning was much slower than in Clayton's original study. Group 4-2, for example, made an average of at least 20 errors in 52 trials in each study as compared to 12.28 for Clayton's 4-2 Ss. Variability was also much greater, the standard error of the mean total errors in 52 trials ranging from 2.35 in Experiment I to 5.11 in Experiment III, compared to 1.26 in Clayton's study. In spite of these differences, the predictions of the model did not deviate reliably from the obtained data in two of three cases. However, it is recognized that when learning is slow and total errors large as in the present experiments, the reciprocal operation called for by the model to estimate J tends to minimize differences between groups. Thus when errors are many, additional errors have little effect on the estimate of J.

Another important consideration appears to be whether or not trapping occurs, or more correctly, whether or not error probability approaches a common asymptote of zero for all Ss. In the present experiments the model fit the data of Experiment II best, where trapping was least. Clayton did not find trapping in his 4-2 Ss. However, in a paper presented by Clayton (1966), trapping was reported for Ss choosing between six and three pellets, as was failure to predict the behavior of this group from the errors of Ss choosing between 6 or 3 and none. It is suggested that the model may not be entirely appropriate when the learning task is such that not all Ss reach the same level of performance.

References

- Clayton, K. N. T-maze choice learning as a joint function of the reward magnitudes for the alternatives. *J. comp. physiol. Psychol.*, 1964, 58, 333-338.
- Clayton, K. N. T-maze choice between unequal reward magnitudes. Paper read at meetings of the Midwestern Psychological Association, Chicago, 1966.
- Spear, N. E., & Hill, W. F. Adjustment to new reward: Simultaneous and successive contrast effects. *J. exp. Psychol.*, 1965, 70, 510-519.
- Spear, N. E., & Spitzner, J. H. Simultaneous and successive contrast effects of reward magnitude in selective learning. *Psychol. Monogr.*, 1966, 80, No. 10 (Whole No. 618).

Note

1. This research was supported by a Predoctoral Fellowship and Grant MH-08888 from the National Institute of Mental Health.