

# Sequential concept formation, hypothesis testing, and the PRE effect

R. A. YAROUGH

*University of Denver, 2030 South York, Denver, Colorado 80210*

High school subjects received training on a sequential concept for monetary reward, were then exposed to partial reinforcement (PR), and finally were extinguished. Percent reward and number of N-R transitions were varied in the PR stage. The effects of PR on subject's hypotheses about the correct pattern were examined. Results of this study indicate that a PRE effect does exist within the constraints of this task. Partial reinforcement was also found to affect the length, variability, and complexity of subject's hypothesis. Contrary to expectations, these effects were independent of number of N-R transitions.

Learning researchers have examined a variety of interpretations of the partial reinforcement extinction (PRE) effect, viz, that variations in percent reward produce differential resistance to extinction. Foremost among these researchers is Capaldi (1967), who introduced the "sequential variable hypothesis" as fundamental to the PRE effect. Capaldi and associates (Capaldi, 1967; Capaldi & Kassover, 1970; Meyers & Capaldi, 1970) have found that the structure of the acquisition reward and nonreward trial sequence has a substantial effect on extinction in animals. In Capaldi's terms, the acquisition variables, N-length (number of successive nonrewards preceding a reward trial) and number of N-R transitions (change from a nonreward or N trial to a reward, R, trial) influence extinction behavior in a direct function regardless of percent reward and are the fundamental factors underlying the typical PRE.

The sequential model has been extended to human extinction behavior in a PR gambling task by Halpern and associates (Halpern & Poon, 1971; Meyers, Driessen, & Halpern, 1972; Poon & Halpern, 1971). These researchers have found that persistence in betting during extinction, i.e., a period of no payoff, is also influenced by manipulations of N-length and number of N-R transitions. Halpern and Poon (1971) have suggested that the effects of these sequential variables are related to the hypothesis or pattern testing activities of human subjects. Subjects confronted by a low percent reward and a particular reward/nonreward structure, e.g., changes in number of N-R transitions, appear to develop patterns or hypotheses of different number and kind than continuous reward (CR) control subjects, and they

are more likely to maintain these patterns into extinction (Halpern & Poon, 1971). Their basic notion, then, is that variations in resistance to extinction may come about as a joint effect of sequential variables and pattern testing tendencies.

The present experiment was undertaken to determine directly whether subjects' acquisition response patterns are affected by nonreward trial sequences (i.e., changes in number of N-R transitions) and whether, in turn, these affect extinction behavior. To explore subjects' response patterns, a blank trials procedure, termed pattern probe, was incorporated into the basic PR gambling design developed by Halpern and Poon (1971). Moreover, while we used the quasigambling format of Halpern and Poon, we also required each subject to learn a sequential concept for correct betting responses to an arbitrary criterion at the outset. It is assumed that learning a response pattern will eliminate pattern *testing* because the subject will have a pattern that "works" and thereby provide a means of controlling pre-experimental biases in pattern or hypothesis testing.

## METHOD

### Subjects

The subjects were 90 high school student volunteers with a mean age of 17 years 2 months. Ten subjects were randomly assigned to each of the nine conditions of the experiment, seven experimental and two control.

### Equipment

The equipment was the same as used by Halpern and Poon (1971). Briefly, the subject sat at a console containing four major units, among which were a bet and a response panel and coin acceptance and payoff boxes. The subject was supplied with a predetermined number of nickels. With the onset of the "bet ready" light, the subject made a decision to bet or not. To bet, he deposited a nickel; a no-bet decision carried a one-cent fine. After 5 sec, the response ready light came on, and subject made a position response by pressing a switch for either the right or left outcome light. If correct, the subject won \$.10; if incorrect, the subject lost his nickel. Seven consecutive no-bets concluded the experiment.

This paper is sponsored by Joseph Halpern, University of Denver. It is based on a thesis submitted in partial fulfillment of the requirements of the M. A. degree at the University of Denver, 1973. The author wishes to thank Joseph Halpern for his advice and support of this project. Additional acknowledgments are gratefully extended to L. E. Bourne, Jr. and to Research Grants MH-14314 from the National Institute of Mental Health and GB-340-77X from the National Science Foundation.

### Task and Design

The present experiment involves three main phases, plus two pattern probes in the seven experimental conditions only. Probe A came at the outset of the experiment, each experimental subject having to respond with a pattern for 10 trials without reinforcement. Phase 1, in which the subject had to learn a sequential concept, right-right-left, immediately followed. During Phase 1, the subject was rewarded on those trials in which he produced the sequentially correct response, i.e., left after two rights. The criterion of learning was nine consecutively correct responses.

After reaching this criterion, the subject transferred to Phase 2, the formal PR acquisition phase of 30 trials. The groups divided at this point on the basis of percent reward and number of N-R transitions. Six experimental groups were formed by the factorial combinations of three reward percentages (60, 45, or 30) and two levels of N-R transitions (2 or 6). A seventh experimental group received 100% reward. Two control groups received the extremes of reward and number of N-R transitions (100%-0 transitions and 30-6). These control subjects did not receive the pattern probes but had an equal-time break after PR acquisition.

With the completion of Phase 2, the experimental groups responded once again to a 10-trial pattern probe (B), producing the pattern they thought was correct, and would maximize their winnings. During extinction, Phase 3, which began at the end of Probe B or the corresponding break, no subject received positive reinforcement for any response he made, i.e., no response was now correct though the subject was allowed to continue to bet and respond. The experiment ended when the subject refused to bet for seven consecutive trials.

## RESULTS AND DISCUSSION

Four measures were analyzed: number of trials to acquisition criterion, number of trials to extinction criterion, characteristics of patterns obtained through the blank trial probes, and number of trials during PR acquisition on which subjects deviated from the original right-right-left pattern (deviation score). Tests for acquisition and extinction performance differences between control groups (those not receiving the pattern probes) and their experimental counterparts were all nonsignificant,  $t_s(18) \approx 1.0$ .

### Initial Performance and Acquisition

No differences among the seven experimental groups were expected in the performance of the first pattern probe. The data were scored for variations in *length* (number of responses in each unique and unrepeated response pattern), *variability* (number of response transitions in a pattern of given length) and *complexity* (product of length  $\times$  transitions). A 3 by 2 ANOVA (percent reward vs. number of N-R transitions as variables) indicated no statistically significant differences, all  $F_s < .88$ . The subjects' predictive behavior and kinds of response patterns generated were, thus, assumed to be unaffected by initial biases in sampling of subjects. Mean between-group differences in the number of trials needed to reach three consecutive repetitions of the right-right-left pattern during Phase 1 were small and nonsignificant, all  $F_s < 1.59$ .

### Partial Reinforcement

Under PR, the subject is informed at various times (on N trials) that his choice, an element of a pattern, is incorrect. It is reasonable to assume that he will attempt to discover the *currently* correct pattern, thereby deviating from the originally learned pattern (Halpern & Poon, 1971; Meyers, Driessen, & Halpern, 1972). Thus, his deviation score should increase to a level of chance responding as the number of "incorrect" (no payoff) trials increases, that is, with decreases in percent reward. Those subjects who receive CR during this phase (information on every trial indicating that the original pattern is still correct) should deviate from that pattern (if at all) less often than subjects who encounter some number of trials of no payoff.

The subject's responses during the PR acquisition phase (Phase 2) were converted to deviation scores by comparing patterns in Phase 2 of length 3 to the originally learned pattern (right-right-left) in search of concurrence. All Phase 2 trials were searched for the pattern, were marked off when it occurred, and remaining trials were counted as deviations. To test for differences among groups in deviation score, a 3 by 2 analysis of variance was conducted with percent reward and number of N-R transitions as the variables. Only the main effect of the percent reward was reliable,  $F(2,54) = 3.76$ ,  $p < .03$ . In an overall one-way analysis, the CR and the three PR groups differed significantly in number of deviations,  $F(3,76) = 132.73$ ,  $p < .001$ .

Deviation scores generally increased as percentage of reward decreased, as can be seen in Table 1. This indicates that subjects who experienced more nonreward trials are more likely to search for a new pattern. Alternatively, the higher the percent reward, the more likely that subject will not deviate and will retain the same or a close approximation of the pattern learned in Phase 1.

### Performance Probe B

Partial reinforcement should enhance hypothesis testing, whereas continuous reinforcement should depress it (Halpern & Poon, 1971). Thus, differences in the kind of hypotheses given by subjects during pattern probe trials can be expected as a result of Phase 2 training. Hypothesis length, variability, and complexity measures of the probe patterns were expected to increase as percent reward decreased. Subjects who received low percent reward were expected to give longer, more complex, patterns than subjects at higher levels of reward and especially different from patterns produced by CR subjects.

The second pattern probe, Probe B, was, like the first, scored for length, variability, and complexity. Analyses of variance identical to those used on the first probe indicated that the only significant effect lies in the percent reward variable,  $F_L(3,76) = 9.10$ ,  $p < .01$ ;  $F_V(3,76) = 4.70$ ,  $p < .01$ ;  $F_C(3,76) = 4.53$ ,  $p < .01$ .

**Table 1**  
Means and Standard Deviations of Phase 2 Deviation Scores, Probe B, and Extinction

	Deviation Scores Phase 2		Length		Probe B Variability		Complexity		Trials to Extinction	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	30 - 2	8.00	.67	6.53	2.83	1.96	1.57	15.59	17.00	49.70
30 - 6	7.70	.83	6.20	2.78	2.80	1.60	21.40	17.20	45.20	24.67
2 + 6	7.85	.74	6.36	2.89	2.38	1.69	18.50	17.80	47.45	26.83
45 - 2	7.10	1.20	7.70	2.36	3.00	2.19	18.86	23.38	38.00	28.89
45 - 6	6.70	1.34	6.40	2.97	2.80	1.66	22.60	18.82	39.40	20.33
2 + 6	6.90	1.25	7.05	2.84	2.90	2.00	20.73	17.89	38.70	24.32
60 - 2	7.20	1.55	6.40	2.80	3.00	1.81	24.20	18.98	20.90	10.70
60 - 6	7.10	.99	7.10	2.11	2.10	.70	15.50	8.34	29.80	29.43
2 + 6	7.15	1.27	6.75	2.57	2.55	2.96	19.85	11.71	25.35	22.03
CR	1.25	1.37	3.43	1.19	1.20	.60	4.83	5.39	22.55	12.98

Note—The format, e.g., “30 - 2,” refers to percentage reinforcement (30) and number of N-R Transitions (2). The values of Probe B in the CR condition are derived from the experimental group; the remaining CR figures are from the combined experimental and control groups.

Pattern length and variability generally increased as percent reward decreased. As seen in Table 1, subjects in the CR condition produced the shortest, least complex, patterns, which were typically identical to the original right-right-left pattern. The 45% and the 60% groups gave more variable and more complex patterns than either the 30% or the CR groups. The difference between continuous and partial reward groups was expected. The complex trend among PR groups was not. Subjects rewarded on 30% of their trials may experience so low an amount of reward that every pattern tried appears to fail. When the subject is asked to predict the pattern that will occur on the next 10 trials (Probe B), he regresses to a modified version of the pattern he learned, thus producing only a slightly more variable response sequence. Alternatively, subjects in the 45% and 60% groups may experience what appears to be a near chance amount of reward (50%). Payoff information confirms a variety of patterns, but none completely, leaving the subject to search for the additional feature which will produce 100% payoff. These subjects would then be generating patterns of greater complexity and length than 30% subjects during the probe.

Overall, the general predictions of a pattern testing model of the PRE (Halpern & Poon, 1971) tended to be supported by the data. Partial reward had some obvious effect on response patterns generated and tested by subjects. Patterns given by the PR subjects were variable and quite discrepant from the originally learned pattern. CR subjects gave stereotypic patterns nearly identical to the response sequence of the first phase, with very few deviations from it.

**Extinction**

The extinction data demonstrated a reliable PRE effect. The number of trials to extinction criterion was strongly related to percent reward, the number of trials increasing as percent reward decreased. Relevant means

are shown in Table 1. The 3 by 2 analysis of variance yielded a significant effect of percent reward,  $F(2,54) = 3.98, p < .03$ , as did the one-way analysis of variance using the CR data,  $F(3,76) = 5.53, p < .05$ . Number of N-R transitions had no separate reliable effect.

Performance during extinction was not affected by subjects' standing above or below the median, during acquisition,  $t(18) \approx 1.00$ . Contrary to typical findings (Grossen, 1969; Halpern & Poon, 1971; Lewis & Duncan, 1958), subjects who took longer to reach criterion during acquisition were not necessarily those who took longer to extinguish. We can attribute this lack of relationship to the unique sequential concept learning feature of the study. In effect, the subject did not learn a single response per trial but a rule which covered three, and possibly more, trials. The extinction measure typical of this type of learning applies to individual trial responses and not subject's perseveration on a sequential concept. Trials to acquire a successful rule or concept is probably a completely different process than acquiring a particular discrete response and might not be expected to covary with measures of individual response strength such as trials to extinction.

**Sequential Variables**

Contrary to several other studies (e.g., Poon & Halpern, 1971), there was no relationship between number of N-R transitions and trials to extinction, the PRE effect, in the present study.

N-R transitions may not be a useful concept for dealing with extinction effects based in the kind of sequential learning involved here. There are several possible reasons. First, N-length, the number of successive nonrewards which precede a reward trial, was not controlled in this experiment. As has been noted before, N-lengths, N-R transitions, and percent reward are impossible to separate completely in a study of partial reinforcement (Halpern & Poon, 1971; Poon &

Halpern, 1971). It may be that a sequential variable is responsible for the PRE effect, as Capaldi (1967) has suggested, but that, in this case, the variable is N-length and not N-R transitions.

It might also be noted that the paradigm used in this experiment was atypical of PRE research. Generally, a two-phased procedure (acquisition and extinction) is used. The present paradigm involved, in addition, a unique sequential concept learning phase. The subjects learned a sequential concept before proceeding to the PR acquisition and extinction phases. Concept learning may negate any effect of the subsequent sequential variables on extinction behavior.

Not only did number of N-R transitions fail to influence trials to extinction, it also had no effect on any measure of response pattern characteristics. The only significant differences in length, variability, and complexity of hypotheses resulted from changes in percent reward. That difference was primarily attributable to the distinction between continuous and partial reward. Thus, there is no support in these data for the notion that an increase in the number of N-R transitions will increase resistance to extinction by increasing the variability and complexity of patterns tested by subjects (Halpern & Poon, 1971).

## REFERENCES

- Capaldi, E. J. A sequential hypothesis of instrumental learning. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1). New York: Academic Press, 1967.
- Capaldi, E. J., & Kassever, K. Sequence, number of nonrewards, anticipation and intertrial interval in extinction. *Journal of Experimental Psychology*, 1970, 84, 470-476.
- Halpern, J. & Poon, L. Human partial reinforcement extinction effects: An information processing development from Capaldi's sequential theory. *Journal of Experimental Psychology Monograph*, 1971, 89, 207-227.
- Meyers, L. S., & Capaldi, E. J. Resistance to extinction as a function of sequence of events in partial reinforcement. *Psychonomic Science*, 1970, 19, 199-200.
- Meyers, L. S., Driessen, E., & Halpern, J. Transfer following regular and irregular sequences of events in a guessing situation. *Journal of Experimental Psychology*, 1972, 92, 182-190.
- Poon, L., & Halpern, J. A small trials PREE with adult humans: Resistance to extinction as a function of number of N-R transitions. *Journal of Experimental Psychology*, 1971, 91, 124-128.
- Rubenstein, I. Some factors in probability matching. *Journal of Experimental Psychology*, 1959, 57, 413-416.
- Spivey, J. E. Resistance to extinction as a function of number of N-R transitions. *Journal of Experimental Psychology*, 1967, 75, 43-48.

(Received for publication April 28, 1975.)