DOCUMENT RESUME

ED 058 590 CG 006 836

AUTHOR Dockstader, Steven L.
TITLE The Effects of Partial Reinforcement in the Acquisition and Extinction of Recurrent Serial Patterns.
INSTITUTION Denver Univ., Colo.
PUB DATE May 71
NOTE 24p.; Paper presented at Rocky Mountain Psychological Association convention, Denver, Colo., May 12-15, 1971

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Discrimination Learning; *Experimental Psychology; *Extinction (Psychology); Feedback; *Learning; Reinforcement; *Sequential Learning

ABSTRACT

The purpose of these 2 experiments was to determine whether sequential response pattern behavior is affected by partial reinforcement in the same way as other behavior systems. The first experiment investigated the partial reinforcement extinction effects (PREE) in a sequential concept learning task where subjects were required to learn a recurrent binary sequence up to a criterion and then were shifted to an extinction phase. The 2nd experiment was conducted to replicate the essential findings of the first study. Several conclusions are drawn from the results: (1) response perseverance in sequential learning tasks may be due to the proactive effects of the informative feedback, not necessarily to the consequences of the event; (2) a PREE in sequential learning tasks results from the proactive effects when discriminative cues for forthcoming, positive feedback are established; and (3) extinction in these more complex learning tasks may not be related to loss of habit strength, but rather to the mere abandonment of one strategy in search of another. (TA)

The Effects of Partial Reinforcement in the Acquisition and Extinction of Recurrent Serial Patterns

Steven L. Dockstader
University of Denver

Recent attempts at descriptions of human choice behavior have tended to emphasize the organizational components of the behavior. In general, the theories tend to characterize S as an active information processor who encodes, presumably for the purpose of later recovery, series of prior decisions (responses) as well as the outcomes of these decisions (correct, incorrect). These characterizations typify the sequential choice models of Restle and Brown (1971) as well as Myers (1971). Interestingly, the characterizations are also consonant with Capaldi's (1967) sequential theory of animal instrumental conditioning.

Much of the impetus for the information processing view of human choice behavior derives from experiments which have used uncertain outcome prediction tasks. In several experiments where trial outcomes were contingent on responses and Ss were probabilistically correct independent of their responses, e.g., Yellot (1969), Halpern and Poon (1971), it was clear that Ss generate their own characteristic trial outcome structures which are evident from specific response patterns that are of sufficient strength to transfer through various forms of extinction training. In fact, many experiments have capitalized on this clear proclivity of Ss to form sequential solutions in uncertain outcome prediction situations. Typical of these are the run length studies, e.g., Restle (1966), where homogeneous series of trial outcomes occur according to some probabilistically predictable outcome schedule and Ss can maximize correct responding by learning the run contingencies.

The present experiments also focus on this tendency toward patterned responding but from a somewhat different point of view. In the experiments reported here, the pattern is treated as a singular unit of behavior on the assumption that it reflects a central behavioral process. That is, since the data suggest that Ss approach uncertain outcome situations as problem solving tasks where the solution presumable involves discovering the "correct" sequential outcome structure, it seems more appropriate to investigate the acquisition of response patterns rather than the singular units of choice behavior. The Ss in the present studies were required to learn a specific outcome pattern of right-left predictions via a modified Humphrey's board where feedback was given by means of lights corresponding to each response button. To a considerable extent then, the experimental context was more similar to a concept learning or serial anticipation task than to traditional binary choice tasks. In fact, Bruner, Wallach and Galanter reported on virtually an identical task in this Journal considering the learning of recurrent regularity.

The purpose of the present experiments was to determine whether sequential response pattern behavior is affected by partial reinforcement in the same way as other behavior systems. While the partial reinforcement extinction effect (PREE) has received considerable attention in a wide variety of experimental contexts, the phenomenon has received only limited and scattered attention in concept learning and problem solving situations. The reason for the lack of attention may, in part, be a function of the fact that reinforcement in such tasks is considered to be primarily informative, serving only a steering or guiding function (Bourne, 1966). However, consideration of the strength of reinforcement effects in other experimental contexts, e.g., choice behavior, where it also serves a steering or guiding function, suggests reinforcement to be a potentially important factor in concept learning.

The limited amount of available research dealing with the PREE in concept learning has yielded conflicting results. Erlebach and Archer (1961) varied percentage of reinforcement and degree of learning in a card sorting task. In a situational analog to extinction, the initially learned correct concept was changed along a single dimension. Performance during the post-shift phase would be expected to include a larger number of perseverative errors as evidence of the existence of a PREE. The expected inverse relationship between number of perseverative errors during the post-shift phase and percentage of reinforcement was not found. The design was, however, confounded by the inclusion of a degree of original learning variable (number of pre-shift trials), and O'Connell (1965) suggested that the effect of this was to increase the variability in the number of reinforcements between groups. O'Connell controlled for the absolute number of reinforcements between groups and found a clear PREE as measured by the number of post-shift perseverative errors between groups receiving 60% and 100% reinforcement. A later experiment by O'Connell and Wagner (1967) was designed specifically to study the PREE in concept learning. They eliminated the transfer analogy in favor of a true extinction period where all responses received negative feedback. A PREE was demonstrated for groups having the same reinforcement schedule as in the earlier study. They also found extinction performance to be independent of variation in number of training trials in opposition to the findings of Erlebacher and Archer (1961).

It is interesting to note that all of the above studies at least allude to the importance of the patterns of reinforcement. A good deal of more recent research in widely disparate areas has emphasized the importance of patterns of reinforcement in situations where reinforcement is in the form of informational feedback (cf Restle & Brown, 1971). These investigations have shown that the pattern of reinforcement is an important component of any learning task. In fact, it may be that Ss attempt to learn the reinforcement pattern in much the same way that they attempt to learn concepts, and a relevant component for learning, which later influences transfer performance, may be the S-defined reinforcement pattern. The available research on such pattern learning is scarce. Bruner et al. (1959) sought to identify the factors that influence the learning of a fixed recurrent sequence of events. The S's task was to predict, on an element by element basis, the structure of the outcomes of a sequence of binary events while being initially unaware of the sequence or its length. In many respects, this is quite similar to a verbal learning task using the method of serial anticipation. Recurrent regularity, as it was referred to by Bruner et al., has subsequently received some attention in attempts at computer simulation of cognitive behavior (Gregg, 1967)

The first experiment investigated the PREE in a sequential concept learning task where Ss were required to learn a recurrent binary sequence up to a criterion and then were shifted to an extinction phase.

Method

Subjects.--The Ss were Denver University introductory psychology students. A total of 49 Ss were randomly assigned to one of three acquisition conditions resulting in 15 Ss per group with four Ss not meeting an acquisition criterion.

Apparatus.--Each S was seated before a raised panel which was functionally equivalent to a Humphrey's board. A warning light at the top of the panel signalled the beginning of a trial and remained lighted during the 5 sec response interval. Below the response interval light were two red event lights, the occurrence of which S was to predict by pressing one of two contact switches located at the base of the panel. Above and corresponding to the choice switches were two green lights which were illuminated coincident with S's choice.

The pair of red event lights was labeled "program indicators," the green lights "choice indicators," and the warning light at the top "response interval light." To S's left was located a box labeled "sequence programmer" which contained relay switches and an audible solenoid that sounded during each trial, but not, of course, contingent upon S's particular choice. To S's right was located a coin box which delivered a nickel at the end of each correctly predicted sequence.

Control of the conditioning apparatus was programmed through two Hunter timers which allowed 5 sec for the prediction of each event, i.e., the response inteval and 5 sec for the intertrial interval. The reinforcing events and the delivery of the monetary reward were controlled from an adjacent room.

Procedure.--Each S was run individually and was instructed that he was in a problem solving task. The task involved predicting which of two lights would terminate a trial and that he was able to win money if he made a number of consecutive correct predictions. No mention was made of a fixed recurrent sequence. There were two conditions of continuous reinforcement which differed only in the degree of training. In the first, CR_1, S was required to perform the task to a criterion of seven correct iterations of the binary sequence. In the second, CR_2, the criterion was 20 correct iterations. The partially reinforced group, PR, received 50% reinforcement according to a variable ratio reinforcement schedule described below. Continuously reinforced Ss were presented a fixed recurrent sequence that was not contingent upon their responses. The trial outcomes were repeated iterations of 1-0-0 or right-left-left (RLL). If the three elements were correctly predicted, then a nickel was delivered in the coin box. Pilot studies revealed that the effects of the coin were, as far as could be determined, exclusively of informational rather than incentive value--the primary function appeared to be one of signalling the end of the sequence. Thus, S received element by element informative feedback for each separate prediction via the lights and a delineation of the length of the sequence by the delivery of the coin.

In group CR_1, S was required to predict the sequence to a criterion of seven consecutive correct iterations. A pilot study revealed that subsequent to seven correct iterations responding was errorless. However, because of questions relating to the effects of degree of original learning on extinction (cf. Erlebacher & Archer, 1961), group CR_2 received an extended acquisition to a criterion of 20 correct iterations. In addition to this consideration, the CR_2 group provided the logical control for the PR group which recieved approximately 20 reinforced sequences.

Traditionally, partial reinforcement in the binary prediction task has involved negative feedback, i.e., the disconfirmation of a prediction. In those tasks the reinforcement schedule was essentially random, each event being independent of the prior event. However, in the present task where there are sequential dependencies it was not possible to partially reinforce each element of the pattern. Previous research (cf. Bruner _et al_., 1959) and pilot studies revealed that discontinuities in the response pattern would result in search behavior: _S_ would, in effect, adopt another response strategy, seemingly seeking a new sequence. This is another manifestation of the apparent proclivity by _Ss_ to establish a patterned solution (see Halpern & Poon, 1971). It was necessary, therefore, to "shape" _S_'s response in such a way that it would be possible to reliably elicit it under conditions of partial reinforcement. One method of doing this is to use a form of a variable ratio schedule of reinforcement. The PR _Ss_ were treated as the CR _Ss_ until they had produced five correct iterations. At this point the next three predictions were disconfirmed. Then the fixed recurrent sequence was resumed until S produced four iterations, following which the next three responses were disconfirmed. This was continued for three reinforced sequences and one nonreinforced, two reinforced and one nonreinforced, etc., until the initial five to one reinforcement ratio was reversed to one reinforced to five nonreinforced sequences. The effectiveness of this procedure was marked for most Ss, none of which were able to discern its seemingly obvious structure. The procedure resulted in 19 reinforced trials and a variable number of nonreinforced trials. This latter variation occurred because some Ss would become "out of phase" with the sequences during the rather extended periods of nonreinforcement in the latter half of the schedule. Rarely, however, did an S ever receive less than 35% total reinforcement during acquisition. It is noteworthy that this procedure was considerably more effective in establishing response evocation than the use of a random schedule of reinforcement. This may be of no small consequence when one seeks to explain the relative paucity of findings concerned with PREE in human learning situations.

Extinction for all groups consisted of transfer to a situation in which a number of disconfirmed predictions were followed by a confirmation of the first two elements of the concept learned in acquisition and a disconfirmation of the final element. That is, at the beginning of extinction, the first three predictions were disconfirmed. The very next prediction of "R" was confirmed, and this was most frequently followed by an "L" which was always confirmed and another "L" which was disconfirmed. After a minimum of three more trials the first "R" was confirmed and so on until a criterion of four consecutive confirmed "R" did not result in the completion of the sequence. That is, four consecutively confirmed RL responses had to be followed with an R, which would occur less than 7% of the time by chance. The task is thus defined as an extinction condition rather than transfer to a new learning task since the third element is never reinforced and thus nothing new to be learned. The discontinuation of the monetary "reward" further underscored the extinction phase.

Results and Discussion

Acquisition.--There were no significant differences for the three groups in trials to initial acquisition of the concept, i.e., the first time it was produced, the means being 17.29, 12.0 and 13.68 for CR_1, CR_2, and PR, respectively. There were, however, four *Ss* rejected on the basis of an acquisition criterion of 50 trials. It was found in pilot studies that most *Ss* who exceeded fifty trials to learn the pattern either did not learn the sequence or spent an inordinate number of trials doing so. This is consistent with the findings of Bruner *et al*. (1959) who found wide individual variation in learning to predict binary sequences. Although the mean number of discrete trials to initial acquisition were approximately the same, the three groups did differ some with regard to learning. The traditional means of graphic representation, i.e., by trial blocks, is perhaps misleading in this learning task. That procedure would allude to a much greater disparity between the groups that actually existed. A more accurate representation makes use of conditional statistics. What is learned in a sequential learning task is the sequential dependency of the binary (in this case) elements. Figure 1 presents the learning of this task in terms of the conditional probability of predicting the third element in the sequence when the first two elements have been confirmed, i.e., $P(L|RL)$.

Insert Figure 1 about here

Two things are immediately apparent from an examination of Figure 1. First, acquisition proceeds rapidly in this task once the sequence has initially been discovered. Granted that the chance level of $P(L|RL)$ is .50, Ss nonetheless reach near perfect performance levels quite rapidly. Second, although there is a small effect of partial reinforcement, the performance levels for the PR group reflects the efficiency of the random ratio schedule with regard to response elicitation.

Extinction.--The extinction criterion is such that an S is considered extinguished for the learned response if he has been given confirmation of the first two elements of the sequence and failed to produce the final element four consecutive times. More than 25% of the CR_1 group produced exactly that few, while none in the PR group did so. The means for these two groups were 9.27 and 18.38 respectively, which was significant ($t(28) = 3.13$, $p < .01$). The mean for group CR_2 was 10.25, sufficiently similar to that of CR_1 that no significance test was deemed necessary.

The findings demonstrate without equivocation the pervasiveness of the PREE. However, this finding is almost singularly unique in the concept formation/problem solving literature. Of additional interest is the finding of no differences between CR groups regarding extinction performance. Results reported by Grant and Berg (1948) and Grant and Cost (1954) would indicate that under continuous reinforcement the number of perseverative errors is influenced to a considerable degree by the number of reinforcements in acquisition. However, Erlebacher and Archer (1961) found no differences in different criterional groups following 100% reinforcement. The present results are clearly in line with the findings of Erlebacher and Archer.

EXPERIMENT II

In an effort to replicate the essential findings of the first study and to explore further the mechanisms operative in partial reinforcement a second study was designed with a somewhat different extinction procedure.

Method

The stimuli and procedures were the same here as in the previous study, except where noted. The Ss were arbitrarily assigned to either a PR or a CR group, 15 in each. The acquisition procedure was the same, but the extinction procedure differed in two ways. First, instead of terminating the experiment at the extinction criterion, extinction was prolonged for 120 discrete trials during which only the first element (R) of the sequence was confirmed 20 times. Again, if the PREE was operative, the PR group should produce more instances of the remainder of the sequence than should the CR group. It was felt that this would be a more critical test of the effect because the disconfirmation of the second element should signal an incorrect prediction for the third element and result in its occurrence at only a chance level if an information hypothesis is held. The second difference in the extinction procedure was more an extension than a change. Instead of discontinuing extinction trials when the criterion was met extinction was maintained until all Ss had at least 20 opportunities to produce the concept. This is to say that the "R" prediction was confirmed 20 times for each S during the extinction period. The reason for this change in procedure was to examine the course of extinction and S's strategies during a less limited period.

Results and Discussion

Acquisition performance for the two groups in the present experiment were virtually identical to the comparable groups of Experiment I. Table 1 gives the means and other descriptive statistics for the number of trials to the extinction criterion for each group of Experiments I and II according to the Experiment I extinction criterion. As might be expected, the means to criterion in Experiment II were lower than in Experiment I, but nonetheless, the CR and PR groups were significantly different, $t(28) = 3.25$, $p < .01$, providing replicative support for the PREE of Experiment I.

Insert Table 1 about here

Another indication of the strength of the PREE might be found by disregarding the criterion and looking at the extended extinction period. Figure 2, which gives the mean probability of LL given R, presents the extinction performance of the two groups with the preshift asymptotes indicated. It can be seen from the figure that the total number of RLL sequences predicted by the PR group was higher than that of the CR group. The means were 5.73 and 3.44, respectively, which were significantly different, $t(18) = 2.12$, $p < .05$.

Insert Figure 2 about here

Interestingly, across trials and Ss, both groups tended to utilize the information provided by the disconfirmation of the second element in the sequence. That this disconfirmation should signal an incorrect prediction for the entire sequence and result in chance responding for the third element was previously indicated. Indeed, given 300 disconfirmed L predictions (Ss X opportunities/Ss), the CR group subsequently followed with 157 L predictions and the PR group with 156, both clearly at chance level. The previously noted differences then must be explained in terms of the second element only, the first being a constant for both groups. In the CR group, 100 L predictions followed the initially confirmed R, while in the PR group 170 L predictions followed R. If chance level is again 150, then what is as important as the persevera- tive effects of the PREE in accounting for the differences is the depression in the CR group which reflects an apparent reluctance to perform the learned response. This is most not- able when the absolute magnitude of the difference of the groups from chance level is considered.

These findings can most easily be interpreted in terms of a generalization decrement hypothesis of the PREE (Capaldi, 1967). This is most evident as it relates to the performance decrement of the CR groups as just previously discussed. An examination of the response strategies of the Ss further sub- stantiates this view. The typical response of Ss at their initial experience with nonreinforcement, which would be during acquisition for the PR group and extinction for the CR group, was the same. However, at extinction for the PR group the response was different. Of the total of 45 CR Ss tested in both experiments, the initial response to extinction was three consecutive right predictions (RRR) for 38 of them. Similarly, during acquisition, the initial nonreinforced trial for PR Ss resulted in RRR for 27 of the 30 Ss. By contrast, only five of the PR Ss responded with RRR predictions to the initial extinction sequence. Clearly, then, it can be seen that CR Ss could more readily discriminate the onset of extinction.

The performance decrement of CR Ss does not wholly account for the differences, however. Perseveration of the learned response, or what might be called the facilitative effects of partial reinforcement, also contributes to the differences in extinction performance. The training received by PR Ss might be characterized as that of having successive acquisition and extinction conditions during the preshift phase which generalize across the transfer to extinction (e.g., Capaldi, 1967). This viewpoint alludes to the notion that the perseverative effects of partial reinforcement can be accounted for in terms of a discrimination or generalization decrement hypothesis and an examination of the task reveals that this is a likely explanation. Additionally, however, the PREE findings in the present experiments may somewhat be due to the proactive effects of the information in the task. This is to say that once S has determined the sequential dependencies involved he is no longer dependent upon the consequences of each individual prediction, i.e., confirmation of the first element in the pattern not only confirms the single prediction, but also during the course of learning, signals what the next event is to be. This proactive effect of informational feedback partially accounts for the perseveration of responding in a transfer condition. In addition to the proactive effects, which all Ss experience, the groups that demonstrated PREE learned a refinement upon proactive reinforcement. For them acquisition included several experiences with fractional extinction, i.e., nonreward. These periods of nonreward occasioned the opportunity for S to test other sequential hypotheses. That these hypotheses were never supported eliminated them as response alternatives. Then, when reinforcement was again reinstated, it was in the same pattern as previously. These repeated reinforcement and extinction conditions established a discriminative control for a confirmed prediction and strengthened its proactive effect. In effect, a confirmed prediction became a discriminative stimulus for further confirmation--an expectancy developed. The course of the development of this expectancy was demonstrated during acquisition by the increase in the probability of predicting the sequence, as can be seen in Fig. 1. The strength of this cue, or discriminative stimulus, was determined with the onset of extinction and it resulted in the PREE. The Ss in the CR condition had no experience with nonreward and their rather immediate response decrement probably should not be viewed as extinction in the sense of a loss of habit strength, but more in terms of S seeking for new hypothetical sequences, just as PR Ss did during acquisitional nonreward. This obviates the traditional view of extinction and proposes simply that Ss respond to the information in the task, pursuing no longer what has been negated, but instead searching for new patterns of response.

Several conclusions can be made from the present experiments. First, response perseveration in sequential learning tasks may largely be due to the proactive effects of the informative feedback and not necessarily to the consequences of the event, although it is conceded that the consequences originally establish the sequential dependencies. Secondly, a PREE in sequential learning tasks results from these proactive effects when discriminative cues for forthcoming positive feedback are established. This explanation appeals to the strength of the cue as the elicitor of the sequential responses which are responsible for perseveration during an extinction period. A third conclusion regarding extinction appears to be warranted. The present experiments strongly suggest that extinction in these more complex learning tasks may not be related to loss of habit strength, but rather to the mere abandonment of one strategy in search of another. One prominent view of extinction behavior conceives of it in terms of a transfer phenomenon (Capaldi, 1967). This is entirely consistent with the present findings in that response decrement was continually found to be a function of the generalization of information from one condition to another.

REFERENCES

Annett, J. Feedback and human behavior. Baltimore: Penguin Books, 1969.

Bourne, L. E. Human conceptual behavior. Boston: Allyn and Bacon, 1966.

Bruner, J. S., Wallach, M. A. & Galanter, E. H. The identification of recurrent regularity. American Journal of Psychology, 1959, 72, 200-209.

Capaldi, E. J. A sequential hypothesis oi instrumental learning. In K. W. Spence & J. T. Spence (Eds.) The Psychology of learning and motivation, New York: Academic Press, 1967.

Dockstader, S. L. & Halpern, J. The partial reinforcement effect in a concept learning task. Paper presented at the meetings of the Rocky Mountain Psychological Association, Denver, May, 1971.

Erlebacher, A. & Archer, E. J. Perseveration as a function of degree of learning and percentage of reinforcement in card sorting. Journal of Experimental Psychology, 1961, 62, 510-517.

Grant, D. A. & Berg, E. A. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. Journal of Experimental Psychology, 1948, 38, 404-411.

Grant, D. A. & Cost, J. R. Continuities and discontinuities in conceptual behavior in a card sorting problem. Journal of General Psychology, 1954, 50, 237-244.

Gregg, L. Internal representations of sequential concepts. In B. Kleinmuntz (Ed.) Concepts and the structure of memory. New York: Wiley, 1967.

Halpern, J. & Poon, L. Human partial reinforcement extinction effects: An information processing development from Capaldi's sequential theory. Journal of Experimental Psychology, 1971, 89, 207-227.

Myers, J. L. Sequential choice behavior. In G. H. Bower (Ed.) The psychology of learning and motivation. New York & London: Academic Press, 1970.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1966.

O'Connell, D. C. Concept learning and verbal control under partial reinforcement and subsequent reversal and non-reversal shifts. Journal of Experimental Psychology, 1965, 69, 144-151.

Restle, F. Run structure and probability learning: Disproof of Restle's model. Journal of Experimental Psychology, 1966, 72, 382-389.

Restle, F. & Brown, E. Organization of serial pattern learning. In G. H. Bower (Ed.) The psychology of learning and motivation. New York & London: Academic Press, 1970.

Yellot, J. I. Probability learning with non-contingent success. Journal of Mathematical Psychology, 1969, 6, 541-575.

Table 1

I ı Trials to Extinction Criterion for Each Group

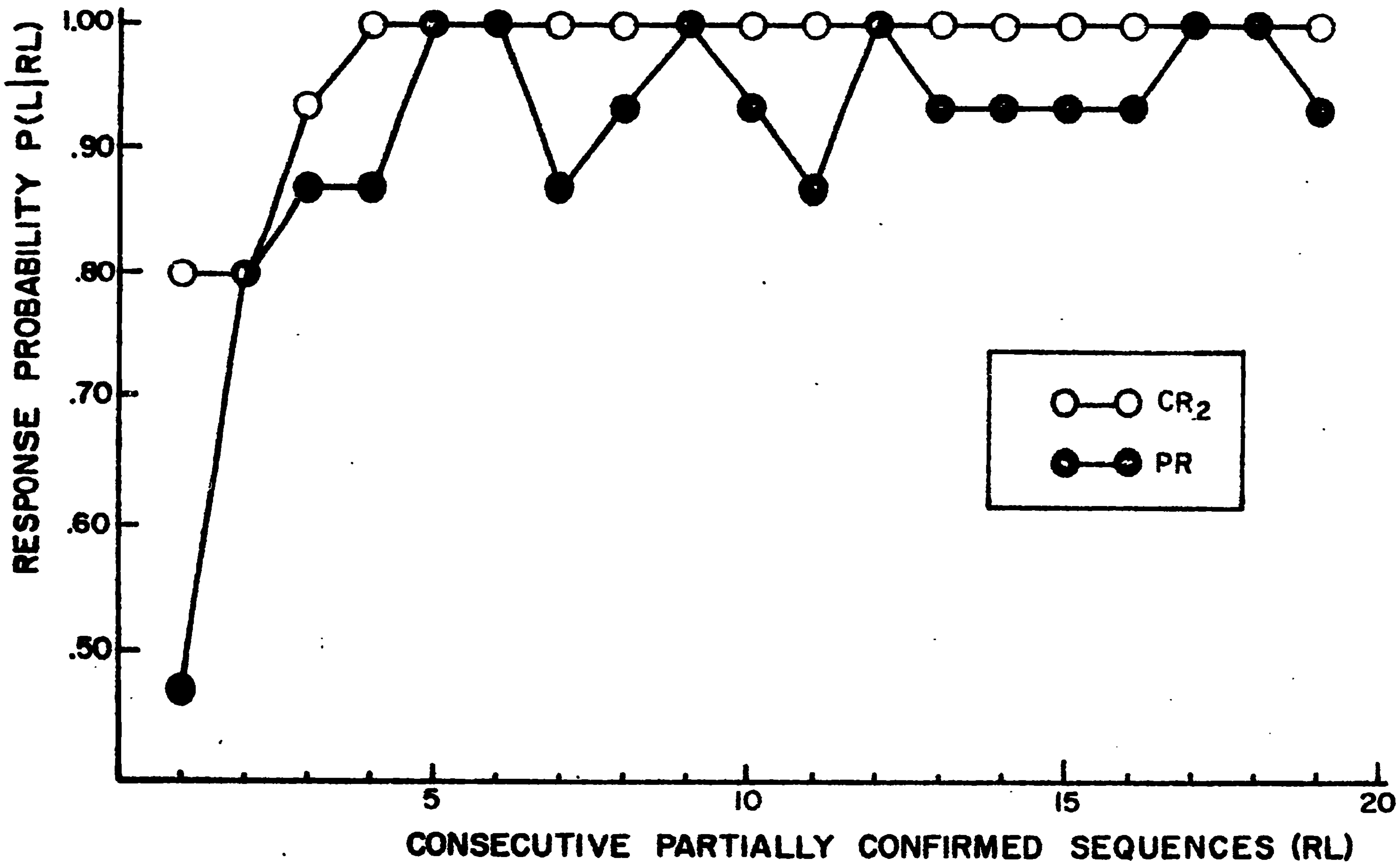
of Experiments I & II

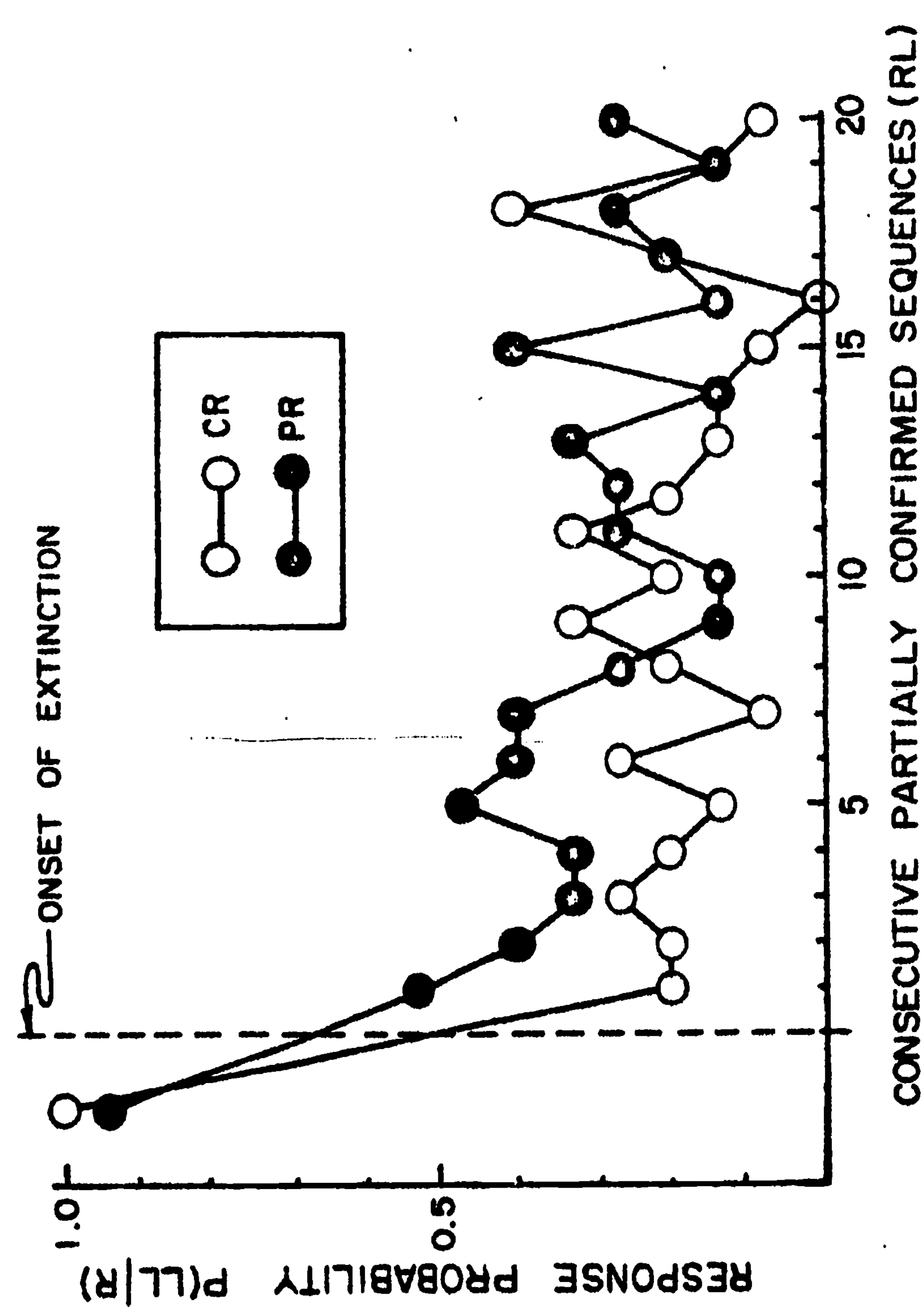
Exp.	Group	Mean	S. D.	Range
I	CR_1	9.27	8.75	4.0 - 22.0
	CR_2	10.25	6.62	4.0 - 23.0
	PR	18.38	7.48	8.0 - 24.0
II	CR	6.27	2.37	4.0 - 13.0
	PR	10.67	4.67	4.0 - 20.0

FIGURE CAPTIONS

Figure 1: Mean probability of a "Left" given confirmation of "Right-Left" on the preceding two trials for each group over all acquistion trials (Exp. I).

Figure 2: Mean probability of "Left-Left" given a confirmed "Right" on the preceding trial over all extinction trials (Exp. II).


RESPONSE PROBABILITY P(L|RL)
1.00
.90
.80
.70
.60
.50
5
10
15
20
CONSECUTIVE PARTIALLY CONFIRMED SEQUENCES (RL)
CR2
PR


ONSET OF EXTINCTION
CR
PR
RESPONSE PROBABILITY P(LL|R)
1.0
0.5
CONSECUTIVE PARTIALLY CONFIRMED SEQUENCES (RL)
5
10
15
20