

Extinction facilitates acquisition of the higher order operant

PAUL T. P. WONG

Trent University, Peterborough, Ontario, Canada K9J 7B8

Higher order operants modify an imposed reinforcement schedule that controls the main operant and render the schedule more favorable. With rats leverpressing for fixed-ratio food reinforcement, an interpolated extinction session facilitated acquisition of the higher order operant, as evidenced in an increase in the appropriate collateral leverpressing which reduced ratio requirements.

Research on frustrative nonreward in partial reinforcement or extinction conditions has centered largely on its invigorating, aversive, and aggressive effects (see Gray, 1971). Several studies have also demonstrated that nonreward increases response variation (Antonitis, 1951; Boroczi & Nakamura, 1964; McCray & Harper, 1962; Newberry, 1971; Wong, 1976), but this effect has not been systematically investigated to the same extent. The main purpose of the present study was to determine whether extinction-induced response variation would serve an adaptive function of reducing the imposed fixed-ratio requirement.

It has been demonstrated that, given a choice of two alternative reinforcing conditions (e.g., ratio or interval values), pigeons will peck a switching or choice key to produce the more favorable (lower ratio or lower interval values) condition and a change in the discriminative stimulus (Findley, 1958; Verhave, 1963). Presumably, the change in the discriminative stimulus functions as a secondary reinforcer for the switching behavior. In the present experiment, rats were given the opportunity to modify a fixed-ratio requirement imposed on the main lever by choosing between two collateral levers and performing on the one that could adjust the imposed schedule and render it more favorable. This type of adjusting operant is called a higher order operant (Wong, 1975) because it modifies and controls schedules of reinforcement that, in turn, modify and control the main operant. Such behavioral control of schedules goes beyond the choice between two given alternative schedules, as in previous research on switching behavior. In the higher order operant paradigm, the subject typically has a wide range of options, and it is possible to achieve various degrees of modification of the imposed schedule in both favorable and unfavorable directions. The subject could, through performing the

appropriate adjusting operant, determine his optimal or preferred reinforcing condition. Skinner (Note 1) considered it as "the second stage in a self-controlling sequence" and stated that "there is a special characteristic in the reinforcement of the second stage resulting from the changed consequences of the first stage." This special characteristic lies in the fact that higher order operants are not maintained by direct consequences (i.e., production of primary or secondary reinforcers) but are *induced* by certain schedules of reinforcement and then maintained by *derived* consequences (i.e., comparison of the previously imposed schedule with the adjusted schedule). The derived reinforcement is always indirect and delayed, because it is derived only after the subject has performed the higher order operant and the main operant to meet the requirement of the new adjusted schedule. A higher order operant is increased and maintained only when it renders the imposed schedule more favorable in terms of reinforcement density or gain/cost ratio. Because higher order operants are defined by a special class of reinforcement that is derived and indirect, they are different from multiple operants (Findley, 1962) and concurrent operants (Catania, 1966).

It has been recently reported that barpressing may be reinforced by a temporary change from a higher probability to lower probability shock (Herrnstein & Heline, 1966) and a change from a low probability of noncontingent food delivery to a more frequent free food delivery (Coulson, 1970). In both cases, the operant was not necessarily reinforced immediately. This type of two-tier arrangement is referred to as a differential probability of reinforcement (dpr) schedule (Keehn, 1972) and may be considered as a special case of higher order operant, in that the reinforcement is derived from a favorable change in schedules of noncontingent events (i.e., events that are not contingent on performing the main operant).

A preliminary study on higher order operants (Wong, Finnie, & Grife, Note 2) has shown that, while most human subjects learned to perform the appropriate adjusting lever to reduce fixed-ratio (FR) requirements, rats showed a strong tendency to fixate on the main operant and failed to learn the higher order operant.

Supported by NRC Grant A8635 and a University Research Committee grant. Thanks are due to Lynn Finnie and Wanda Grife for their assistance in collecting the data. The author is grateful to Evalyn F. Segal, J. D. Keehn, and Rod Wong for their helpful comments on an earlier version of this paper. Requests for reprints should be addressed to Dr. Paul T. P. Wong, Department of Psychology, Trent University, Peterborough, Ontario, K9J 7B8, Canada.

Subsequent study (Wong, Note 3) revealed that extending the training duration only increases the tendency to fixate on the main operant. In the present study, it was predicted that an interpolated extinction session would induce more response variation (e.g., more collateral leverpressing) and facilitate acquisition of the higher order operant.

METHOD

Apparatus

The testing chamber, as illustrated in Figure 1, was made of Plexiglas and measured 33 x 33 x 33 cm. Three 5.2-cm-wide Gerbrands-type aluminum levers were mounted on one side of the wall, 4 cm apart and 4 cm above the grid floor. BRS solid state modules were used for programming and recording in conjunction with a cumulative event recorder.

Subjects

The subjects were 21 light-hooded rats of the Walker-Walker strain bred in the animal colony of Trent University. They were housed in single cages, handled, and placed in a food deprivation regimen when they were approximately 100 days old. The experiment began when subjects were reduced to 80% of their individual predeprivation body weights.

Procedure

All subjects were randomly assigned to one of three conditions: extinction, reinforcement, and control. Rats were tested in squads of three subjects, one from each condition. For successive squads, subjects within each squad were rotated in the order of testing. For example, the running order of Squad 1 was extinction, reinforcement, control; the running order in Squad 2 was reinforcement, control, extinction; the running order in Squad 3 was control, extinction, reinforcement.

On Day 1, all subjects underwent five phases of training: (1) operant level, (2) magazine training, (3) shaping, (4) baseline, and (5) progressive ratio training. Subjects were first placed in the testing chamber individually for 1 h of operant-level test to determine preference between the left and right collateral levers. To insure a reasonable level of responding on these levers, food crumbs and pellets were placed behind the Plexiglas wall on which the levers were mounted. Immediately following the operant-level session, subjects were magazine trained by 30 noncontingent reinforcements (45-mg pellets). Each delivery of a food pellet into the foodcup was preceded by a magazine click.

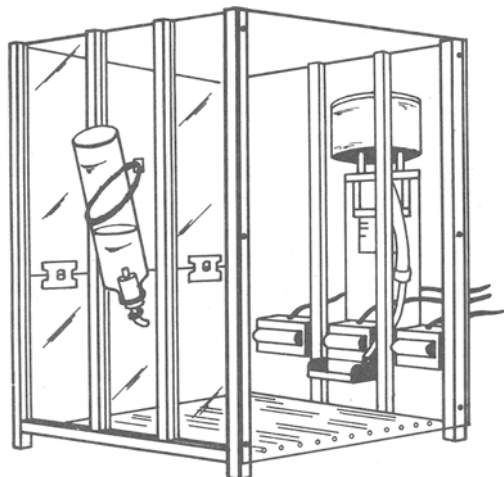


Figure 1. A three-lever conditioning chamber. (The water bottle is optional.)

Subjects were then shaped to press the center (main) lever to the criterion of four consecutive leverpresses and were reinforced for each press (i.e., FR1). Shaping was followed by 5 min of baseline measure during which the food dispenser was disconnected, so that pressing the center lever resulted in neither food reinforcement nor magazine click. The baseline session was intended to further monitor position preference in collateral leverpressing. Finally, subjects were given progressive ratio training in the main operant with incremental steps of six responses per step until the ratio was increased to FR60. Thus, subjects were first given 30 reinforcements at FR1, and then 5 reinforcements following each incremental step in responding on the center lever. Throughout these phases on Day 1, both the left and right collateral levers had no programmed consequences.

On Day 2, all subjects first received one session of steady state training during which both collateral levers remained ineffective, and 40 reinforcements could be received by pressing the center lever on FR60. After the initial session, Group Control received another steady state session, while Groups Extinction and Reinforcement received one session of higher order operant testing, during which the preferred and nonpreferred collateral levers had different programmed consequences. Any two processes on the nonpreferred lever reduced the imposed ratio requirement (FR60) for the main operant by 10 responses, but once the ratio was reduced to FR10, no further reduction was possible. Any single press on the preferred lever reset all responses to zero and reinstated FR60 for the main operant. Therefore, responding on the preferred side reduced reinforcement density and increased gain/cost ratio, while performance on the nonpreferred side had the opposite effect of increasing reinforcement density and reducing gain/cost ratio.

Whenever the imposed ratio requirement (FR60) was reduced by pressing the nonpreferred collateral lever, the subject was allowed a maximum of only 10 reinforcements at the reduced ratio; then the ratio was immediately reset to FR60. Therefore, to maintain a more favorable ratio requirement throughout testing, the subject had to switch to the nonpreferred side repeatedly whenever the ratio requirement was reset to FR60. This higher order operant session lasted until the subject had received 40 reinforcements.

On Day 3, Group Control received two sessions (40 reinforcements per session) of steady state training. Group Reinforcement received two higher order operant sessions, while Group Extinction received one session of extinction treatment followed by one higher order operant session. During extinction, collateral levers remained operative, but the main operant produced only secondary reinforcement in the form of magazine clicks and the dropping of a food pellet into a dish outside the testing chamber. The extinction session lasted until the subject had received 40 secondary reinforcements or until 60 min had elapsed, whichever came first.

RESULTS AND DISCUSSION

Results for the major experimental sessions are summarized in Table 1. During the operant-level session, all subjects except one exhibited a position preference. The one exception exhibited preference during baseline. It is clear from the table that position preference remained in all sessions in conditions where collateral levers were ineffective.

During progressive ratio learning, no switching behavior occurred until FR12. Figure 2 shows that mean number of collateral leverpressings and mean number of changeovers (i.e., switching to collateral levers) progressively increased with ratio size and peaked at FR48, after which a decremental trend set in. Any

Table 1
Mean Frequencies of Collateral Leverpressing on Initially Preferred (P) and Nonpreferred (N) Sides

1		2		3		4		5		6		7	
P	N	P	N	P	N	P	N	P	N	P	N	P	N
Extinction Group													
29.29*	14.14*	1.86**	.86**	22.86†	12.00†	9.57††	4.00††	19.85†*	7.29†*	25.14*†	69.00*†	4.14†*	41.43†*
Reinforcement													
36.29*	25.14*	8.43**	.00**	21.00†	18.29†	12.29††	7.43††	17.00†*	8.29†*	10.71†*	10.00†*	3.29†*	6.86†*
Control													
30.29*	20.57*	1.57**	.31**	17.00†	9.14†	10.86††	5.86††	21.00††	10.00††	22.59††	9.57††	5.10††	4.85††
*Operant Level		**Baseline		†Progressive Ratio		††Steady State		†*Higher Order Operant		*†Extinction			

interpretation of these collateral lever responding data must account for both the incremental and decremental trends. One interpretative possibility is superstitious conditioning. However, this hypothesis would predict only an incremental trend and the tendency for collateral responding to occur prior to food reinforcement. Since all leverpresses were recorded on the cumulative recorder, it was possible to examine the temporal distribution of collateral leverpressing with respect to reinforcement. Such an analysis revealed that only .07% of collateral leverpressing occurred during postreinforcement pauses, and prereinforcement collateral responding distributed approximately equally across different segments of various ratios; for example, the number of changeovers and the number of collateral responses during the first half of FR18 were the same as those during the second half of the ratio. The superstitious conditioning hypothesis also has difficulty accounting for the absence of collateral responding under FR1 and FR6. Since the collateral leverpressing was not maintained by inadvertent reinforcement or secondary reinforcement, it may be regarded as belonging to the same category as schedule-induced adjunctive polydipsia and aggression (Gilbert & Keehn, 1972).

The present finding may be accounted for by the interaction of two opposing response tendencies. As ratio size increases, the tendency of response variation

also increases. However, progressive ratio training necessarily involved extended reinforced practice of the main operant, which is assumed to increase the tendency of response fixation. When the terminal ratio size is not very large (e.g., FR60), the tendency to fixate on the main operant may become stronger than the tendency of response variation. The interaction of these opposing tendencies should produce a curvilinear function in collateral leverpressing as well as in the latency of performing the main operant, as reported by Wong and Amsel (1976).

It should be noted that collateral leverpressing was maintained in Group Control throughout the entire experiment, although both collateral levers had no programmed consequences. In Group Reinforcement, although preferred and nonpreferred collateral levers became effective in Session 5 and had differential effects on the imposed schedule of reinforcement, their collateral responding did not significantly differ from Group Control in Sessions 5, 6, and 7; there was no significant reversal in preference even though the initially nonpreferred side could increase reinforcement density. Thus, corroborating previous findings by the author, there was no evidence of learning the higher order contingency by Group Reinforcement.

In Session 6 only the extinction treatment resulted in a significant increase in collateral leverpressing over the preceding session ($t = 3.92, p < .01$) and produced a reversal in preference with a significantly higher level of responding on the initially nonpreferred side ($t = 4.60, p < .01$). This reversal in preference continued in Session 7 ($t = 3.34, p < .01$) even though the extinction treatment had already terminated. Leverpressing on the initially nonpreferred side was significantly higher than Group Reinforcement ($t = 2.98, p < .05$) and Group Control ($t = 2.94, p < .05$) during Session 6, and the superiority of Group Extinction relative to Groups Reinforcement and Control ($t_s = 2.45, p < .05$ in both cases) extended into Session 7; but during the same two sessions, no significant group differences were obtained with respect to the initially preferred side. Acquisition of the higher order operant by Group Extinction is also evident in reduction of ratio requirements. The mean numbers of responses per reinforcement during the last two sessions for Groups Extinction

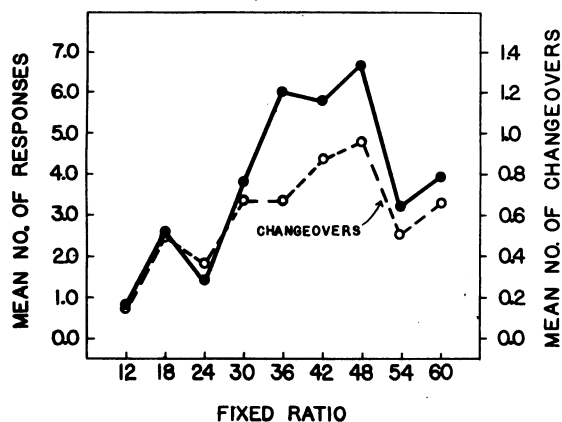


Figure 2. Collateral leverpressing (both left and right levers) during progressive ratio training.

and Reinforcement were 35.9 and 55.9, respectively ($t = 5.17, p < .01$).

Reversal in preference in Group Extinction throughout Sessions 6 and 7 can only be accounted for in terms of reinforcement derived from a favorable comparison of the imposed schedule with the adjusted schedule following responding on the initially nonpreferred side. This derivation process requires at least two memories (one for the imposed schedule value and one for the adjusted schedule value) and the remote association of the more favorable reinforcing condition with the appropriate collateral leverpressing which occurred prior to completing the main operant. Therefore, it is a rather difficult contingency to learn.

The present experiment provides some evidence of learning the higher order contingency when a high level of collateral responding was induced by extinction. Previous research on experimental extinction has demonstrated its effects on inhibition (Amsel, 1971a), regression (Amsel, 1971b), aggression (Gallup, 1965; Knutson, 1970), and aversion (Adelman & Maatsch, 1956; Wagner, 1967). The present experiment demonstrates that the effects of extinction need not be negative, that extinction may increase response variation and promote the learning of a difficult higher order contingency.

The paradigm of higher order operants provides a methodology to explore hitherto unsuspected adaptive potentials of various schedule-induced adjunctive behaviors which appear pathological or toxic in their maladaptive excessiveness (Falk, 1971). It raises the intriguing question of whether schedule-induced toxic behavior (e.g., polydipsia) would develop when animals are given the opportunity to learn an adaptive higher order operant. This paradigm also enables one to assess the efficiency of coping with schedule demands by various species or subpopulations in each species, because it provides a situation where an individual has the option of being controlled by an unfavorable schedule of reinforcement imposed on him or attempting to control the schedule. If he chooses the latter, there are also different degrees of efficiency in controlling the schedule. For example, in the present study, the optimal degree of efficiency is to keep the ratio requirement constantly at FR10 by performing a minimum number of responses on the adjusting lever.

REFERENCE NOTES

1. Skinner, B. F. Personal communication, October 21, 1974.
2. Wong, P. T. P., Finnie, L. M., & Grife, W. *Second-order operant under FR conditions*. Paper presented at the annual meeting of the Psychonomic Society in Boston, Massachusetts, November 1974.
3. Wong, P. T. P. *Higher order operant acquisition under FR conditions with extended training*. Unpublished study, Trent University, 1975.

REFERENCES

- ADELMAN, H. M., & MAATSCH, J. L. Learning and extinction based upon frustration, food reward, and exploratory tendency. *Journal of Experimental Psychology*, 1956, **52**, 311-315.
- AMSEL, A. Positive induction, behavioral contrast, and generalization of inhibition in discrimination learning. In H. H. Kendler & J. Spence (Eds.), *Essays in neobehaviorism*. New York: Appleton-Century-Crofts, 1971. Pp. 217-236. (a)
- AMSEL, A. Frustration, persistence, and regression. In H. D. Kimmel (Ed.) *Experimental psychopathology: Recent research and theory*. New York: Academic Press, 1971. (b)
- ANTONITIS, J. J. Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology*, 1951, **42**, 273-281.
- BOROCZI, G., & NAKAMURA, C. Y. Variability of responding as a measure of the effect of frustration. *Journal of Abnormal and Social Psychology*, 1964, **68**, 342-345.
- CATANIA, A. C. Concurrent operants. In W. J. Honig (Ed.) *Operant behavior: Areas of research and application*. New York: Appleton-Century-Crofts, 1966.
- COULSON, G. E. Positive reinforcement as an increase in the availability of food. Unpublished PhD dissertation, York University, 1970.
- FALK, J. L. The nature and determinants of adjunctive behavior. *Physiology and Behavior*, 1971, **6**, 377-388.
- FINDLEY, J. D. Preference and switching under concurrent scheduling. *Journal of the Experimental Analysis of Behavior*, 1958, **1**, 123-144.
- FINDLEY, J. D. An experimental outline for building and exploring multi-operant behavior repertoires. *Journal of Experimental Analysis of Behavior*, 1962, **5**, 113-166.
- GALLUP, G. G., JR. Aggression in rats as a function of frustrative nonreward in a straight alley. *Psychonomic Science*, 1965, **3**, 99-100.
- GILBERT, R. M., & KEEHN, J. D. *Schedule effects: Drugs, drinking, and aggression*. Toronto: University of Toronto Press, 1972.
- GRAY, J. G. *The psychology of fear and stress*. New York: McGraw-Hill, 1971.
- HERRNSTEIN, R. J., & HINELINE, P. N. Negative reinforcement as shock frequency reduction. *Journal of the Experimental Analysis of Behavior*, 1966, **9**, 421-430.
- KEEHN, J. D. Schedule-dependence, schedule-induction, and the law of effect. In R. M. Gilbert & J. D. Keehn (Eds.), *Schedule effects: Drugs, drinking, and aggression*. Toronto: University of Toronto Press, 1972. Pp. 66-94.
- KNUTSON, J. F. Aggression during the fixed-ratio and extinction components of a multiple schedule of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1970, **13**, 221-232.
- MCCRAY, C. L., & HARPER, R. S. Some relationships of schedules of reinforcement to variability of response. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 19-21.
- NEWBERRY, B. H. Response variability and the partial reinforcement effect. *Journal of Experimental Psychology*, 1971, **89**, 137-141.
- VERHAVE, T. Towards an empirical calculus of reinforcement value. *Journal of the Experimental Analysis of Behavior*, 1963, **6**, 525-536.
- WAGNER, A. R. Frustration and punishment. In R. N. Haber (Ed.) *Current research in motivation*. New York: Holt, Rinehart, & Winston, 1967.
- WONG, P. T. P. The concept of higher order operant: A preliminary analysis. *Bulletin of the Psychonomic Society*, 1975, **5**, 43-44.
- WONG, P. T. P. A behavioral field approach to instrumental learning: I. The partial reinforcement effects and sex differences. *Animal Learning & Behavior*, 1976, in press.
- WONG, P. T. P., & AMSEL, A. Prior fixed-ratio training and durable persistence. *Animal Learning & Behavior*, 1976, **4**, 461-466.