

# Response–outcome associations remain functional through interference treatments

ROBERT A. RESCORLA

*University of Pennsylvania, Philadelphia, Pennsylvania*

Prior reports have shown that instrumental responses are selectively depressed by devaluing the outcomes they have earned. This occurs even when the response–outcome association has been subjected to a decremental procedure, such as extinction or replacement of that outcome by another. From these results it has been concluded that response–outcome associations are unaffected by various decremental procedures. However, since devaluation takes several days, it is possible that the decremental treatments do attenuate the response–outcome association but it then spontaneously recovers during the time of devaluation. Four experiments investigated this possibility by replicating previous findings under circumstances that allow assessment of recovery for other, concurrently treated, behaviors. The primary responses continued to show full sensitivity to devaluation of their outcomes, despite evidence that recovery of the other behaviors was far from complete. This suggests that the original response–outcome association was preserved and fully functional throughout the decremental and devaluation phases of the experiment.

There is now excellent evidence that instrumental training leads to the formation of an association between a response (R) and its consequent outcome (O). Moreover, once that R–O association has been formed, it appears to be remarkably resistant to procedures that one might have expected to be destructive. For instance, after R–O training, the extinction of R, the arranging for it to occur independently of O, or the replacing of O with a different outcome, all appear to leave the original R–O association intact (Rescorla, 1991, 1992, 1993a, 1993b).

These conclusions are based on data that come from two techniques for exposing the R–O association, devaluation and transfer. In the devaluation technique, the R–O association is assessed by the degree to which post-acquisition devaluation of a particular O results in suppression of the R that previously produced that O. In the transfer technique, the R–O association is assessed by the degree to which a stimulus (S) that elsewhere signals a particular O also results in augmentation of the R that previously produced that O. Each of these techniques has found evidence that the R–O association survives a variety of depressive manipulations.

However, each of these techniques has its interpretative difficulties, leading to concern that each might be insensitive to actual attenuation of the R–O association. For instance, the transfer technique involves the presentation of a stimulus that might serve as a powerful retrieval cue for the R–O association. Therefore, it might be successful in exposing even a highly weakened R–O association. The devaluation technique may be insensi-

tive to reduction in the R–O association for a different reason. In a typical devaluation procedure, repeated pairing of O with a toxic agent, such as LiCl, takes place over several days. The result is that several days necessarily intervene between the end of any procedure intended to destroy the R–O association and the testing of the state of that association. That passage of time may provide an opportunity for substantial recovery from the effects of the destructive procedure, resulting in an underestimation of its initial impact on the R–O association.

The present experiments focus on the possible shortcomings of the devaluation procedure. They seek to evaluate whether its detection of a continued R–O association after destructive attempts is attributable to this opportunity for recovery during the course of devaluation. Several incidental results suggest that any recovery that occurs is not so substantial as to undermine seriously the usefulness of the devaluation procedure. In a recent series of experiments, Rescorla (1995) trained a response first with one outcome and then with another. Then either the first-used or second-used outcome was devalued, with the intention of assessing the degree to which devaluation of each undermined performance. He found similar depression of responding whether he devalued the first or the second outcome, suggesting that the two R–O associations make equivalent contributions to performance. But the particular outcomes employed in those experiments happened to yield somewhat different overall response rates. This made it possible to ask whether the response rate in the test, after the devaluation period, reflected the rate controlled by the first- or second-used outcome. The finding of interest was that the test response rate was primarily controlled by the second-used outcome. If recovery of the first-stage learning occurred during devaluation, it was not so substantial as to control overall

---

This research was supported by National Science Foundation Grants BNS-88-03514 and IBN94-04676. Correspondence should be addressed to R. A. Rescorla, Department of Psychology, University of Pennsylvania, 3815 Walnut St., Philadelphia, PA 19104.

performance. Yet the devaluation results suggested that both the R-O associations from both stages contributed equally.

The experiments reported here attempted to assess more explicitly the possibility of recovery of initial learning from any destructive attempts of an intervening treatment. The strategy was to "tag" different stages of the experiment. At the same time that the target response received its initial training, a discrimination was concurrently trained between another pair of responses. Then at the time when the target response was subjected to a decremental procedure, the other pair of responses received reversal training. The result is that one can ask, during the test of the target responses, whether performance of the other pair of (tagging) responses reflects the discrimination trained during initial target response training or the reversal trained at the time the target response was undergoing decrement.

## EXPERIMENT 1

This experiment was intended to explore a procedure for tagging different sources of performance. It examined the destructive effects on an R-O1 association that might result from training that response with another outcome. Previous reports have found such a manipulation to have relatively little impact on the R-O1 association (Rescorla, 1993b). Performance of the response continued to be adversely affected when O1 was devalued, despite the intervening reinforcement of the response with another outcome. This experiment repeated those procedures with the addition of a behavioral tag to allow assessment of the degree to which each phase of training controlled the animal's performance.

Figure 1 illustrates how this was done. Two target responses, R1 and R2, were rewarded by two different outcomes, O1 and O2. The responses were counterbalanced as lever press and chain pull; the outcomes were counterbalanced as a pellet and liquid sucrose. In the same sessions, another pair of responses, R3 and R4 (a nose poke and a handle pull) were given discrimination training with O3 (polycose) as the reward. Then an attempt was made to undermine the R1-O1 and R2-O2 associa-

tions by deleting both O1 and O2 but replacing them with the common O3. During this phase of the experiment, the original discrimination between R3 and R4 was reversed. Next, devaluation was conducted for one of the outcomes by pairing it, over a 10-day period, with LiCl. Finally, the state of the R1-O1 and R2-O2 associations was assessed by looking for the differential impact of the outcome devaluation. During those same test sessions, however, R3 and R4 were also present, to determine whether their performance was primarily controlled by the original discrimination (learned during the R1-O1 and R2-O2 training) or by the reversal (learned during the R1-O3 and R2-O3) training. The question was whether we would see differential sensitivity of R1 and R2 to the treatment of their initial outcomes despite the continued control of R3 and R4 by the training conducted at a time when R1 and R2 were rewarded with the common outcome.

## Method

### Subjects and Apparatus

The subjects were 16 male Sprague-Dawley rats about 90 days old. They were housed in individual cages and maintained on a food deprivation regime that kept them at 80% of their ad lib body weight. They had free access to water in the home cage.

The apparatus consisted of eight operant chambers measuring 22.9 cm × 20.3 cm × 20.3 cm, identical to those used in previous reports (e.g., Colwill & Rescorla, 1985). The two end walls of each chamber were aluminum; the side walls and ceiling were clear Plexiglas. Each chamber had a recessed food magazine in the center of one end wall. Two small metal cups measuring 1.25 cm in diameter and 1.5 cm in depth were sunk side by side in the floor of each food magazine. To the left of the magazine was a lever and to the right was a chain suspended from a microswitch mounted on the lid of the chamber. Located directly above the food magazine was a 2-cm opening behind which was an aluminum plate that activated an attached microswitch when displaced by a nose poke. Inserted under the grid floor just to the right of the magazine aperture was a flat metal rod, one end of which was bent back to form a handle. An upward pull on this handle operated a microswitch. Access to these manipulanda could be blocked by covering the lever with a metal shield, retracting the chain through a hole in the ceiling, covering the nose poke opening with a jeweled lens, and removing the handle pull from under the grid floor. The floor of the chamber was composed of 0.48-cm stainless steel rods, spaced 1.9 cm apart. Each chamber was enclosed in a sound- and light-resistant shell. The outside ceiling of the shell supported two solenoid-operated gravity feed valves that were connected via plastic tubing to the cups in the food magazine. One system permitted the presentation of .3 ml of an 8% sucrose solution; the other permitted the presentation of .3 ml of a 15% polycose solution. Also attached to that food magazine was a dispenser containing 45-mg pellets (P. J. Noyes Co., Formula A).

Experimental events were controlled and recorded automatically by relays and microprocessors located in an adjoining room.

### Procedure

**Initial training.** On each of the first 3 days the animals received a 20-min magazine training session with pellets, sucrose, and polycose, respectively. Each session contained 20 deliveries of one outcome. Over the next 2 days, all animals were trained to lever press and chain pull. Each training session allowed responding to earn 25 deliveries of an outcome on a continuous reinforcement schedule. For half the animals, lever pressing led to

Phase 1	Phase 2	Devaluation	Test
R1-O1	R1-O3	O1-LiCl	R1
R2-O2	R2-O3		R2
R3-O3	R3-		R3
R4-	R4-O3		R4

**Figure 1.** Design of Experiment 1. Two responses (R1 and R2) were first trained with differential outcomes (O1 and O2) and then with a common outcome (O3). One outcome was then paired with LiCl and both responses tested. Two other responses (R3 and R4) received discrimination training and reversal mixed with the treatments of R1 and R2.

a pellet and chain pulling to sucrose; for the other half of the animals, the response–outcome relations were interchanged. On the next 2 days all animals were trained to nose poke and handle pull, using polycose as the outcome. Throughout this initial training, individual shaping was used if necessary for a particular response. On each of the next 2 days, all animals received two 20-min training sessions spaced about 30 min apart. Each session contained one manipulandum on which responding was reinforced on a variable-interval (VI) schedule with a mean of 1 min. On the 1st of these days, lever pressing and chain pulling were rewarded with their appropriate reinforcers. On the 2nd day, nose poking and handle pulling were rewarded with polycose.

**Phase 1 training.** On each of the next 8 days, all animals received two 20-min training sessions, spaced 30 min apart. In each session, both the nose poke and the handle pull manipulanda were available. In each session, half the animals had nose poking rewarded and handle pulling nonrewarded; the other half of the animals had the reinforcement contingencies reversed. Each session also contained either the lever or the chain. When present, that manipulandum was rewarded by the outcome it received in initial training. Reinforcement assignments were such that for half of the animals receiving the lever–pellet and chain–sucrose relations, the nose poke was rewarded and the handle pull was nonrewarded, whereas for the other half, the handle pull was rewarded and the nose poke was nonrewarded. All outcomes were scheduled on independent VI 1-min schedules.

**Phase 2 training.** On each of the next 8 days, all animals received training identical to that of Phase 1, with two exceptions. First, the sucrose and pellet outcomes were both replaced with polycose. Second, the reinforcement contingencies for the nose poke and handle pull were reversed.

**Devaluation.** Next, the animals received five 2-day cycles intended to devalue one of the outcomes. Throughout this period all manipulanda were removed from the chambers and one outcome (either pellets or sucrose) was paired with a 5% body weight .6 M LiCl ip injection. On the 1st day of each cycle, one outcome was programmed for delivery 20 times at a rate of 1/min. At the end of the session the animal was removed from the chamber, given the LiCl injection, returned to its home cage, and fed 2 h later. On the 2nd day of each cycle, the procedure was the same with the other outcome, but no injection was given. On the third cycle, as the animals reduced their intakes, the number of outcomes delivered was reduced to 10; on the fourth and fifth cycle only five deliveries were made. The outcome to be devalued was selected so that for half the animals it was pellets and for half it was sucrose, counterbalanced across the other treatments.

**Test.** On the next day, the animals received two 10-min test sessions. Each session made available (1) either the lever or the chain and (2) both the nose poke and the handle pull. No outcomes were delivered during these sessions.

**Results and Discussion**

Initial training proceeded smoothly. By the end of Phase 1, the mean response rate per minute was 11.7 for the target responses, the chain, and the lever. Although responding for pellets was higher than that for sucrose, there were no reliable differences as a function of manipulandum or outcome. The discrimination between handle pull and nose poke was also quite good. On the last Phase 1 session, the mean responses per minute were 12.8 and .2 for the rewarded and nonrewarded manipulandum, respectively.

The course of performance during Phase 2 is shown in Figure 2, which displays performance on the final day of Phase 1 and throughout Phase 2 for the target lever and

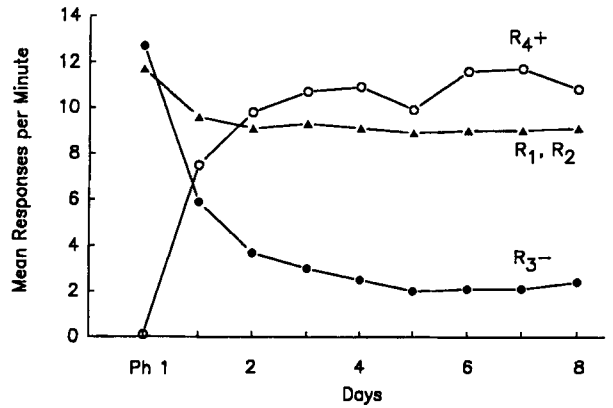


Figure 2. Mean responses per minute to the target responses and to the R3 and R4 responses undergoing reversal. Data are from the final day of Phase 1 and all days of Phase 2.

chain responses as well as for the discriminated nose poke and handle pull responses. With the institution of the new contingencies, the mean responses per minute on the lever and chain (now rewarded by polycose) dropped slightly but remained stable through the remainder of Phase 2. Responding on the nose poke and handle pull rapidly reversed in accordance with the reversed reward contingencies.

Over the course of devaluation, the animals came to reject the outcome paired with LiCl but continued to consume that outcome not paired with LiCl. On the final devaluation cycle, the mean number of pellets left uneaten (of 5 delivered) was 4.3 and 0 for the animals receiving pellets devalued or not, respectively. All animals receiving sucrose devaluation left some in the cups; none of the other animals left any sucrose.

The data of most interest, from the test sessions, are shown in Figure 3, which displays the mean response rates in the extinction test sessions for the target lever and chain responses and for the discrimination handle pull and nose poke responses. The results of responding

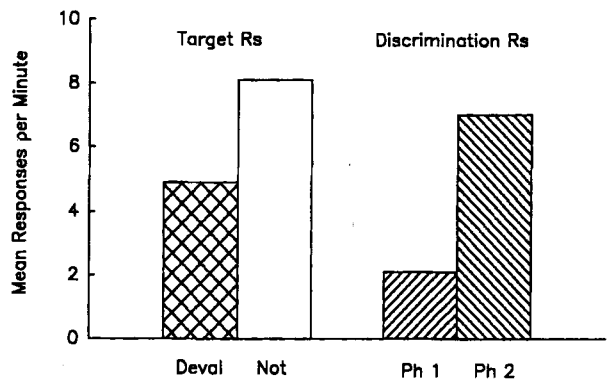


Figure 3. Test results of Experiment 1. One target response had been paired with LiCl (Deval) whereas the other had not. The discrimination responses had been rewarded either in Phase 1 or Phase 2.

on the target responses are shown on the left, separated according to whether the outcome used to train that response in Phase 1 had been devalued or not. It is clear that despite the intervening training of both responses with polycose, responding depended on the treatment of the Phase 1 outcome. The lower level of performance for the response whose Phase 1 outcome had been devalued was reliable [Wilcoxon  $T(16) = 24, p < .05$ ].

The results of discriminative performance on the nose poke and handle pull are shown on the right side of Figure 3. Responding is identified for the response rewarded in Phase 1 (but nonrewarded in Phase 2) and the response rewarded in Phase 2 (but not in Phase 1). It is clear that discriminative performance was dominated by the Phase 2 training. There was reliably more responding on the manipulandum rewarded in Phase 2 but not in Phase 1 [ $T(16) = 7.5, p < .01$ ]. At the same time, discriminative performance was attenuated relative to that prior to devaluation. The discrimination ratios, formed by dividing responding on the reinforced response by total responding, were .89 and .77 at the end of training and after devaluation, respectively [ $T(15) = 3, p < .01$ ].

These results confirm those of earlier experiments in showing that a response remains sensitive to devaluation of its outcome even when training with another outcome intervenes between original training and devaluation (e.g., Rescorla, 1993b). More importantly, that sensitivity can be observed under circumstances in which the other behavior of the animal indicates that the intervening training continued to dominate performance. This suggests that the time needed to conduct devaluation did not allow full recovery of Phase 1 learning; the learning that occurred in Phase 2 continued to control behavior. Consequently, it seems unlikely that the Phase 2 training initially eliminated control of Phase 1 learning only to permit its full recovery during the period of devaluation. Of course, it remains possible that Phase 2 training had some negative impact on Phase 1 learning by the chain and lever.

This experiment suggests that one could use the present procedures to tag the various phases of training and assess their relative current contributions. The subsequent experiments exploit that in order to assess the relative current control over responding by the Phase 1 and Phase 2 outcomes.

## EXPERIMENT 2

As noted, Rescorla (1993b) reported that a response that was first rewarded with one outcome and then with another was equally sensitive to devaluation of either outcome. This suggests that training with a second outcome had no adverse effect at all on the state of the first R-O association. Indeed, it suggests that both R-O associations contribute equally to performance despite the fact that one outcome was used more recently. However, this conclusion is weakened by the possibility that training with the second outcome initially made it ascendent over the first outcome, but that the contribution of the first outcome recovered during the time required for devalu-

ation. This possibility seems especially attractive within the context of some retrieval theories (see, e.g., Bouton, 1993). The recency of the second phase of learning might make it initially dominant. However, with the passage of time, the two phases might become relatively more balanced in their contributions to performance.

The goal of the present experiment was to evaluate this possibility. The design is schematized in Figure 4. In the manner of Experiment 1, two responses (R1 and R2) were initially trained with differential outcomes concurrently with discriminative training of two other responses, R3 and R4. Then, the outcomes used for R1 and R2 were interchanged; at the same time, the discrimination between R3 and R4 was reversed. This means that the R1-O1 and R2-O2 associations were trained in the presence of an R3+/R4- discrimination, but the R1-O2 and R2-O1 associations were subsequently trained in the presence of an R3-/R4+ discrimination. Then O1 was paired with LiCl. In the final test, responding to R1 and R2 was assessed. The question of interest is which response would be more adversely affected, R1 (which had initially earned O1 and then O2) or R2 (which had the order of the reinforcers reversed). This allows us to determine the relative contribution of the first and second earned outcomes during the test. But, in the manner of Experiment 1, the likelihoods of R3 and R4 were also assessed. Just prior to devaluation, of course, R4 should dominate R3. However, if there is substantial recovery over time of the Phase 1 learning, then one would expect responding to R3 and R4 to become similar. Of most importance, if there were full recovery of the Phase 1 contribution, then responding to R3 should dominate that to R4.

## Method

### Subjects and Apparatus

The subjects were 16 male rats of the same sort and maintained in the same manner as those in Experiment 1. The apparatus was that of Experiment 1.

### Procedure

The procedure was identical to that of Experiment 1 with two exceptions. First, initial training of all responses was conducted with polycose prior to Phase 1. This was done to guarantee identical experience for each response with the outcomes given in Phase 1 and Phase 2. Second, during Phase 2 training, instead of

Phase 1	Phase 2	Devaluation	Test
R1-O1	R1-O2	O1-LiCl	R1
R2-O2	R2-O1		R2
R3-O3	R3-		R3
R4-	R4-O3		R4

Figure 4. Design of Experiment 2. Two responses (R1 and R2) were first trained with differential outcomes (O1 and O2) and then with those outcomes interchanged. One outcome was then paired with LiCl and both responses tested. Two other responses (R3 and R4) received discrimination training and reversal with another outcome (O3) mixed with the treatments of R1 and R2.

the lever and chain both being rewarded with polycose, each received the same outcome as the other had received in Phase 1.

### Results and Discussion

Phase 1 performance was like that of Experiment 1. By the final day of training, the mean responses per minute were 9.0 on the target lever and chain responses. The mean responses per minute were 13.8 and 0.2 for the rewarded and nonrewarded discriminated responses, respectively. Over the course of Phase 2, the level of performance on the lever and chain grew somewhat to 12.1 responses per minute. Responding rapidly reversed for the nose poke and handle pull. On the final day of Phase 2, the mean responses per minute were 13.7 and 2.5 for the reinforced and nonreinforced responses, respectively. During the final cycle of devaluation, the animals receiving pellets paired with LiCl left a mean of 3.9 of the 5 pellets but consumed all the sucrose; the animals devalued with sucrose all left sucrose in the wells but consumed all of the delivered pellets.

The data of primary interest, from the test sessions, are shown in Figure 5. The left portion of that figure displays performance on the target responses, separated according to whether the first-earned or second-earned outcome had been devalued. It is clear that performance was equivalent whichever outcome was paired with LiCl. This suggests that each outcome is contributing equally to performance in the test.

The right portion of Figure 5 shows responding on the discriminated responses. As in Experiment 1, performance was dominated by the Phase 2 training. Responding was reliably greater for the response that was rewarded in Phase 2 but not in Phase 1 [ $T(15) = 7, p < .01$ ]. There was, however, an attenuation in the discrimination relative to that observed prior to devaluation. Choice discrimination ratios were .82 at the end of Phase 2 but .72 in the test, a difference that fell just short of conventional levels of reliability [ $T(16) = 31, p <$

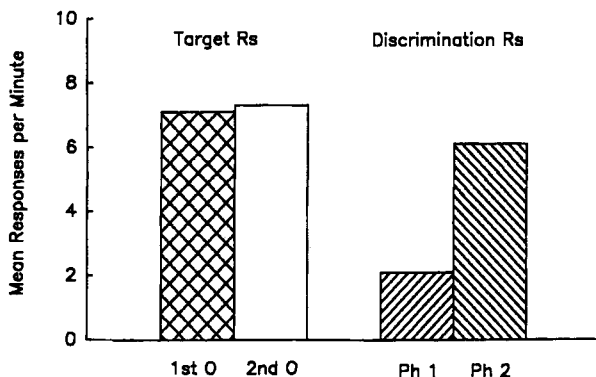


Figure 5. Test results of Experiment 2. The target responses had either their first-earned or second-earned outcome paired with LiCl. The discrimination responses were rewarded either in Phase 1 or Phase 2.

.10]. Consequently, there may have been some recovery of the Phase 1 performance.

These results suggest that the first- and second-earned outcomes contribute equally to the performance of the lever and chain responses. This occurred despite the fact that responding on the nose poke and handle pull indicated that there was far from full recovery of the Phase 1 performance. That observation weakens the argument that the equivalence of control by the two outcomes for the lever and chain is attributable to complete recovery of a previously depressed Phase 1 association.

### EXPERIMENT 3

Another procedure that might be expected to undermine the R-O association is extinction, the opportunity to make the response without any outcome forthcoming. Earlier experiments using the devaluation technique suggest, however, that sensitivity to devaluation of an outcome is unaffected by extinction. In this experiment, those observations were repeated in a context in which the possibility of recovery from extinction is monitored by the method used in Experiments 1 and 2.

The design used in this experiment was identical to that of Experiment 2 except that extinction of both R1 and R2 intervened between Phase 1 and Phase 2 training. As noted elsewhere (e.g., Rescorla, 1991, 1993b), this means that each response is trained with two outcomes; however, there is the opportunity for extinction of the first learned R-O association but not of the second learned R-O association. One can then ask about the relative impact of devaluation of an R-O association subjected to extinction or not.

### Method

#### Subjects and Apparatus

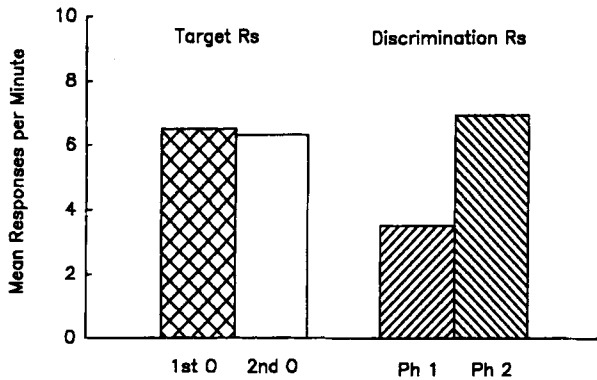
The subjects were 16 male rats of the same sort and maintained in the same manner as those in Experiment 1. The apparatus was that of Experiment 1.

#### Procedure

The procedure was identical to that of Experiment 2 with one exception. Between Phase 1 and Phase 2 training each animal had 5 days during which both lever and chain were subjected to extinction. On each day the animals received two 20-min extinction sessions, one with each response. During this treatment the nose poke and handle pull manipulanda were absent from the chambers.

### Results and Discussion

Phase 1 performance was like that of previous experiments. By the final day of training the mean responses per minute were 10.4 on the target lever and chain responses. The mean responses per minute were 14.0 and 0.4 for the rewarded and nonrewarded discriminated responses, respectively. Over the course of Phase 2 the level of performance on the lever and chain declined somewhat, to 8.8 responses per minute. Responding rapidly reversed for the nose poke and handle pull. On the final day of Phase 2, the mean responses per minute



**Figure 6.** Test results of Experiment 3. The target responses had either their first-earned or second-earned outcome paired with LiCl. The discrimination responses were rewarded either in Phase 1 or Phase 2.

were 13.4 and 3.9 for the reinforced and nonreinforced responses, respectively. Devaluation proceeded as in previous experiments. On the final cycle, the animals consumed all of the outcomes not paired with LiCl, but those animals receiving devaluation of pellets left a mean of 4 of the 5 pellets delivered and those animals receiving devaluation of sucrose all left sucrose in the wells.

The data of primary interest, from the test sessions, are shown in Figure 6. The left portion of that figure displays performance on the target responses, separated according to whether the first-earned (and therefore extinguished) or second-earned (and therefore nonextinguished) outcome had been devalued. It is clear that performance was equivalent whichever outcome was paired with LiCl, suggesting that each outcome is contributing equally to performance in the test, despite the differential exposure to extinction.

The right portion of Figure 6 shows responding on the discriminated responses. As in Experiments 1 and 2, performance was dominated by the Phase 2 training. Responding was reliably greater for the response that was rewarded in Phase 2 but not in Phase 1 [ $T(16) = 24, p < .05$ ]. As in previous experiments, there was an attenuation in the discrimination relative to that observed prior to devaluation. Choice discrimination ratios were .79 at the end of Phase 2 but .68 in the test [ $T(16) = 24, p < .05$ ]. Consequently, there was some reliable recovery of the Phase 1 performance.

These results suggest that even with differential opportunity for extinction, the first- and second-earned outcomes contribute quite similarly to performance. As in previous experiments, responding on the discriminated behaviors was not nearly substantial enough to make plausible an account in terms of full recovery of an otherwise depressed Phase 1 association.

#### EXPERIMENT 4

In this experiment, replication of the results of previous experiments was attempted and several of their shortcomings were addressed.

The strength of Experiments 2 and 3 is that they allow comparison of the associations of the same response with its first- and second-used outcomes. This can be expected to provide an especially sensitive test of the relative strengths of these two associations. However, one weakness of this comparison is that equivalent associative strengths produce a result of no differential sensitivity to devaluation during the test. Unfortunately, that same null result could come about in other ways. For instance, it may be that with the extensive and complex training used here, the devaluation procedure has a generally reduced impact on responding. Although other results have suggested that this technique is highly robust, it would be useful to have a demonstration of its continued power in the present context, in which different phases of the experiment are tagged by other behaviors.

Experiment 4 also addressed another difficulty with all of the previous experiments. In these experiments, the target responses were each trained in one session per day, but the tagging discrimination was operative in both of those sessions. The intention was to arrange as close a temporal relation as possible between the target and tagging behaviors. But the consequence was that the target responses received only half as much training as the tagging responses. It seems quite possible that this additional training would result in less recovery of the Phase 1 behavior in the tagging responses than in the target responses. This might lead to underestimation of the degree of Phase 1 recovery in the target responses.

In Experiment 4 a between-group design was used that corrected both of these difficulties. For this purpose, an elaboration of the design used in Experiment 1 was employed; a comparison group was added to allow assessment of the degree to which the Phase 1 associations survived Phase 2 training. Thus, two groups of rats were given two phases of training with both lever and chain. For one group, the lever and chain produced the differential sucrose and pellet outcomes in the first phase of training but the common polycose outcome during the second phase, in the manner of Experiment 1. For the second group, the lever and chain produced differential outcomes, without subsequent interference. This was accomplished by having them earn the common polycose in the first phase and the differential outcomes in the second phase. Both groups then received devaluation with one of the differential outcomes and were tested with both lever and chain. It was anticipated that devaluing an outcome would selectively depress its response whether that response earned the outcome in the first or second phase, thereby showing the continued effectiveness of the devaluation procedure. The data of interest are magnitudes of the selective devaluation effect when the devalued outcome was used in the first or second phase. On the basis of the results of Experiments 2 and 3, it was anticipated that these selective devaluation effects would be similar in size.

Both groups of animals also received discrimination training and its reversal with the nose poke and handle pull manipulanda during Phases 1 and 2, in the manner of previous experiments. However, to better match the

amount and distribution of experience with the tagging and target responses, each animal received one session per day in which both the target responses, lever and chain, were trained and a second session in which the nose poke and handle pull were discriminated. Training the target and discrimination responses in different sessions separates them somewhat in time during the day, but it allows better matching of the amounts of training.

### Method

#### Subjects and Apparatus

The subjects were 16 male rats of the same sort and maintained in the same manner as those in Experiment 1. The apparatus was that of Experiment 1.

**Initial training.** On each of the first 3 days, the animals received a 20-min magazine training session with pellets, sucrose, and polycose, respectively. Each session contained 20 deliveries of one outcome. Over the next 4 days, all animals were trained to lever press, chain pull, nose poke, and handle pull. Each training session allowed responding to earn 25 deliveries of the polycose outcome on a continuous reinforcement schedule. Throughout this initial training, individual shaping was used if necessary for a particular response. On the next 2 days, the animals received two 20-min training sessions, spaced 30 min apart. One manipulandum was present in each session and responding on it earned polycose on a VI 1-min schedule. Lever pressing and chain pulling were trained on the 1st day; nose poke and handle pull were trained on the 2nd day.

**Phase 1 training.** On each of the next 8 days, all animals received two 20-min training sessions, spaced 30 min apart. In one session both the lever and chain were present and they earned outcomes on a VI 2-min schedule. A 2-min schedule was used for each manipulandum to better match the overall rates of reinforcement for the target and tagging responses. For Group 1, the lever and chain led to the differential sucrose and pellet outcomes during Phase 1, with half the animals having each response-outcome identification. For Group 2, the lever and chain both led to polycose. In the second session of each day, both nose poke and the handle pull manipulanda were available. Half of the animals had nose poking rewarded with polycose on a VI 1-min schedule and handle pulling nonrewarded; the other half of the animals had the reinforcement contingencies reversed. These reinforcement assignments were arranged orthogonally to those for the lever and chain. The order of target and discrimination sessions was varied from day to day in a quasi-random fashion.

**Phase 2 training.** On each of the next 8 days, all animals received training identical to that of Phase 1, with two exceptions. First, the reinforcers used with lever and chain for Groups 1 and 2 were interchanged, so that Group 1 now received polycose for responding on both manipulanda and Group 2 now received the differential sucrose and pellet outcomes. Second, the reinforcement contingencies for the nose poke and handle pull were reversed.

**Devaluation.** This was conducted in a manner identical to that of Experiment 1.

**Test.** On the next day, all animals received two test sessions, each 10 min long. During the first session, the animals had both lever and chain present; during the second session, they had both nose poke and handle pull present. The results of the first test provide information on the sensitivity of the target responses to differential devaluation in Groups 1 and 2. The results of the second test give information on the relative importance of Phase 1 and Phase 2 training of the discrimination.

### Results and Discussion

Phase 1 performance was somewhat lower than that in previous experiments, possibly because of the use of a less

dense VI schedule. By the final day of training the mean responses per minute were 6.3 and 4.0 on the target lever and chain responses, for Groups 1 and 2, respectively. Although it was not reliable, the higher response rate with sucrose and pellets compared with polycose is frequently observed in our laboratory. The mean responses per minute were 8.3 and 0.1 for the rewarded and nonrewarded discriminated responses, respectively. Over the course of Phase 2, the level of performance on the lever and chain declined somewhat to 4.1 per minute for Group 1 and rose somewhat, to 5.8 per minute, for Group 2. Responding rapidly reversed for the nose poke and handle pull. On the final day of Phase 2, the mean responses per minute were 8.6 and 1.5 for the reinforced and nonreinforced responses, respectively.

Devaluation proceeded as in previous experiments. On the final cycle, the animals consumed all of the outcomes not paired with LiCl, but those animals receiving devaluation of pellets left a mean of 4.8 of the 5 pellets delivered and those animals receiving devaluation of sucrose all left sucrose in the wells. There were no differences as a function of instrumental training treatments.

The data of primary interest, from the test sessions, are shown in Figure 7. The left portion of that figure displays performance on the target responses, separated for the two groups and according to whether the differential outcome earned had been devalued or not. Overall performance reflected the response rates at the end of Phase 2, with Group 2 (which had received pellets and sucrose during the second phase) responding at a slightly higher rate. However, for both groups, responding was lower on the manipulandum whose differential outcome had been devalued. This was the case whether that differential outcome had been earned in Phase 1 (for Group 1) [ $T(8) = 2, p < .05$ ] or earned in Phase 2 (for Group 2) [ $T(8) = 4, p < .05$ ]. Moreover, the magnitude of the difference in responding on the two manipulanda was quite similar in the two groups [Mann-Whitney  $U(8,8) = 30$ ]. These observations suggest that the R-O associations learned in Phase 1 are as strong as those learned in Phase 2.

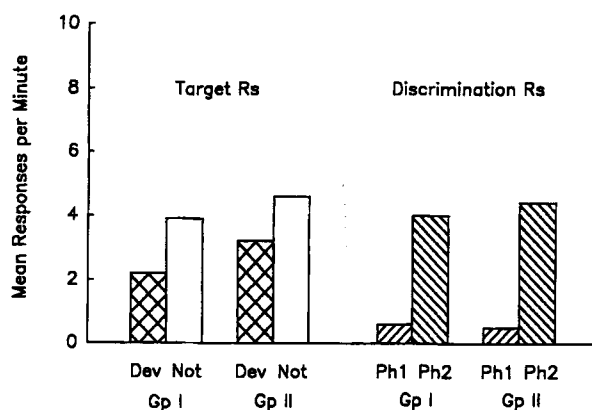


Figure 7. Test results of Experiment 4. The target responses had either their Phase 1 (Group 1) or Phase 2 (Group 2) outcomes devalued or not. The discrimination responses were rewarded in either Phase 1 or Phase 2.

The right portion of Figure 7 shows responding on the discriminated responses, separated for the two groups. As in previous experiments, performance was dominated by the Phase 2 training. Responding was reliably greater for the response that was rewarded in Phase 2 but not in Phase 1. This difference was reliable for each group [ $T_s(8) = 0, ps < .01$ ].

Thus, this experiment confirms the conclusions drawn from Experiments 2 and 3. Despite the dominance of Phase 2 training for performance of the discrimination responses, there was a high degree of similarity in the sensitivity of the target responses to devaluation of their outcomes, whether the first-earned or second-earned outcomes received devaluation. In the present experiment, the two sets of responses were more equivalently treated and the continued effectiveness of devaluation was demonstrated.

### GENERAL DISCUSSION

These experiments confirm earlier reports that R–O associations survive exposure of responding to various potentially destructive procedures. In Experiments 1 and 4, R1–O1 and R2–O2 associations that were well established in Phase 1 survived training of both Rs with a common O3 outcome in Phase 2, as indexed by their differential sensitivity to devaluation of O1 and O2. In Experiment 2, interchanging O1 and O2 during Phase 2, after initial training in Phase 1, resulted in each R being equally susceptible to devaluation of its first- or second-earned outcome. Experiment 3 confirmed this finding even under circumstances in which extinction intervened between training with the two outcomes. Experiment 4 confirmed the results of Experiment 2 under conditions that more closely matched the treatment of the tagging and target responses and that demonstrated the continuing power of the devaluation procedure.

The unique contribution of these experiments is that these observations were made at the same time that the Phase 1 and Phase 2 treatments were tagged by discriminative performance on two concurrently trained responses, R3 and R4. That tagging demonstrated dominant control by the treatments given during the destructive procedures in Phase 2. This in turn suggests that continued control by the Phase 1 R1–O1 and R2–O2 associations is not due to the loss of the learning that occurred during Phase 2 as a result of the time it took to administer devaluation. Rather, those Phase 1 associations seem to have survived the Phase 2 destructive procedures, perhaps without any reduction.

It is important to note that the R3/R4 discrimination established during Phase 2 did show some deterioration over the period it took to devalue the outcomes. Discriminative performance was weaker during the final test compared with its level at the end of Phase 2. A natural interpretation of this result is that there is recovery of reverse R3/R4 discrimination that had been trained in Phase 1, although the absence of a group lacking Phase 1 training prevents drawing that conclusion with certainty.

But it is quite clear that recovery of that Phase 1 R3/R4 discrimination falls far short of being complete.

Of course, it is possible that the recovery of the R1/R2 learning is more complete than is the recovery of the R3/R4 learning. Bouton and Brooks (1993) have noted that on occasion recovery of learning about different stimuli can occur relatively independently. But taken together with the routine observation that recovery is incomplete, the present data strongly suggest the possibility that the Phase 1 learning survives relatively unscathed.

We have interpreted the equivalence of strength in the R–O associations developed in Phases 1 and 2 as evidence that the Phase 1 association is unaffected by Phase 2 training. But these results are also consistent with the possibility that Phase 2 undermines the Phase 1 learning at the same time as Phase 1 has an equivalent adverse effect on Phase 2 learning. Although it is not possible to rule out this possibility on the basis of the present results, data from other experiments make it seem less likely. For instance, Rescorla (1993a) used a transfer design to assess the strength of the association between a response and an outcome when that response either had or had not received further training with a different outcome. He found no evidence that training with the second outcome adversely affected the initial R–O association.

If the response–outcome associations learned in Phase 1 are not adversely affected by Phase 2 learning, then some other basis must be found for phenomena like spontaneous recovery and the partial recovery of Phase 1 R3/R4 discrimination seen here. One possibility, suggested elsewhere, is that many destructive procedures involve the superimposition on the original R–O associations of some response-specific but outcome-independent learning. For instance, extinction might involve the learning not to make a specific response while leaving the R–O associations in place. That might take a variety of forms. One possibility is that an inhibitory association develops between stimuli present and the response itself. Alternatively, it might be that responding in the absence of the original outcome results in frustration (F) that becomes associated with the response. Such an R–F association might come under the control of stimuli present during extinction. Moreover, it might develop even under conditions in which one outcome is substituted for another. Recent results from this laboratory have provided evidence consistent with each of these possibilities (Rescorla, 1993a, 1996).

In any case, these experiments add to the growing literature suggesting that extinction and related procedures leave original associations intact. Moreover, they show the usefulness of a new procedure for assessing those associations at a time when the continuing impact of various treatments can be verified.

### REFERENCES

- BOUTON, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114, 80-99.
- BOUTON, M. E., & BROOKS, D. C. (1993). Time and context effects on



- performance in a Pavlovian discrimination reversal. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 165-179.
- COLWILL, R. M., & RESCORLA, R. A. (1985). Post-conditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 120-132.
- RESCORLA, R. A. (1991). Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 465-474.
- RESCORLA, R. A. (1992). Response-independent outcome presentation can leave instrumental R-O associations intact. *Animal Learning & Behavior*, **20**, 104-111.
- RESCORLA, R. A. (1993a). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, **21**, 327-336.
- RESCORLA, R. A. (1993b). Preservation of response-outcome associations through extinction. *Animal Learning & Behavior*, **21**, 238-245.
- RESCORLA, R. A. (1995). Full preservation of a response-outcome association through training with a second outcome. *Quarterly Journal of Experimental Psychology*, **48B**, 252-261.
- RESCORLA, R. A. (1996). Spontaneous recovery after training with multiple outcomes. *Animal Learning & Behavior*, **24**, 11-18.

(Manuscript received March 8, 1995;  
revision accepted for publication November 28, 1995.)