Transfer of instrumental control mediated by a devalued outcome

ROBERT A. RESCORLA
University of Pennsylvania, Philadelphia, Pennsylvania

In three instrumental learning experiments, rat subjects were used to explore transfer of an instrumental discriminative stimulus to a new response. The stimulus increased the likelihood of that response to the degree that the stimulus and response had a history of association with the same outcome. Moreover, devaluation of the outcome by pairing with lithium chloride had no detrimental effects on its ability to mediate transfer to a new response. This result helps one choose among various accounts of transfer.

An instrumental discriminative stimulus, during which one response is reinforced, comes to control not only that response, but also other responses trained with the same reinforcer. For instance, Colwill and Rescorla (1988) trained rats to nosepoke by using a pellet reward during one stimulus and a sucrose reward during another stimulus. They then separately trained two target responses, leverpressing and chainpulling, by using a pellet reward for one and a sucrose reward for the other. In a subsequent test, they found that the stimuli trained with nosepoke augmented leverpressing and chainpulling to the degree that the stimulus and response shared an outcome. This finding suggests that both the stimulus (S) and the response (R) become associated with the rewarding outcome (O). Consequently, such outcome-based transfer has proven extremely useful for the evaluation of the current state of the R–O and S–O associations (e.g., Rescorla, 1991, 1992).

However, there remains considerable uncertainty about the detailed manner in which this outcome-based transfer acts. The intention of the experiments reported here is to provide evidence on this issue.

One historically attractive possibility has been suggested by two-process theories (e.g., Trapold & Overmier, 1972). According to this view, during discrimination training S develops a Pavlovian association with O. As a consequence, the presentation of S results in the anticipation of O. Since R is then rewarded in the presence of that anticipation, an O–R association also develops. A similar pattern of associations forms in simple nondiscriminated training, except that contextual stimuli play the role of S. As a result of these S–O and O–R associations, the O mediates the ability of S to elicit R. Moreover, the S can transfer its control across responses trained with the same outcome.

This account has much to recommend it. It uses conventional associative mechanisms not only to account for transfer but also to predict and explain a variety of other phenomena (see, e.g., Peterson, Wheeler, & Armstrong, 1978). However, several recent experiments suggest that its reliance on an O–R, rather than an R–O, association, makes the standard two-process account inadequate. For instance, Rescorla (1992) arranged for a transfer response to be rewarded with one outcome (O1) at a time when the stimulus environment led to the anticipation of another outcome (O2). A subsequent transfer test found that response to be augmented by an S otherwise associated with O1 rather than with O2. This result is congruent with others that suggest that instrumental training of R by O has as its primary consequence an R–O, rather than an O–R, association.

Two alternative ways have been suggested in which the R–O association may be involved in outcome-mediated transfer. One possibility is that the R–O association is used in the "backward" direction, so that the presentation of S activates a representation of O, which in turn activates R (e.g., Asratyan, 1974). In its most primitive form, this view is a simple adaptation of the two-process theory, employing the R–O association in the backward direction, rather than the O–R association in the forward direction. A somewhat more complex version of the same account is that S employs the S–O and R–O associations to activate a representation of R. That activation then results in a response to the degree that R is associated with a valuable outcome.

An alternative use of the R–O association derives from the Pavlovian literature on modulation (e.g., Holland, 1983; Rescorla, 1985). According to that view, the response is made to the degree that R activates a representation of a valuable O. The role of the stimulus is to modulate that activation, effectively making it easier for R to activate O. This might be accomplished by modulating either the strength of the R–O association or the overall accessibility of O. In effect, S shifts the threshold for R's activation of O.

Although these conceptions of transfer are quite different, they make generally similar predictions for many situations. But they do appear to differ in the importance that they place on the value of O at the time it mediates responding. According to the threshold-shifting view, S makes it easier for R to activate the mediating O. But one would expect that to promote responding only to the degree that O is a currently valuable outcome. Increasing the access to an O that has lost its value would not be expected to promote responding. However, according to the view that the R-O association is used in the backward direction, the current value of the mediating O should be less important. The main role of O is to activate a representation of R; then R either will occur directly or will be made because of its association with some valuable outcome. But there is no need for the mediating outcome itself to be valuable.

Some evidence is consistent with the possibility that transfer can be mediated by an outcome that has little value. For instance, Rescorla (1990) found evidence for transfer of a stimulus to a new response on the basis of their sharing a neutral feature of the outcome. In this case, the mediator itself was not the valuable outcome that served as the primary basis for responding. Moreover, Colwill and Rescorla (1990) found that a stimulus continued to transfer to a response despite the devaluation of their shared outcome by its pairing with LiCl.

In the present experiments, this possibility was explored more systematically. Experiment 1 was intended to replicate the observation that outcome devaluation does not destroy transfer. Experiments 2 and 3 were attempts to evaluate the magnitude of transfer as a function of the value of the mediating outcome. All of the experiments used conventional instrumental training procedures in which rat subjects made leverpresses and chainpulls for positive outcomes such as solid pellets and liquid sucrose.

EXPERIMENT 1

This experiment was intended to confirm the prior observation that outcome-based transfer continues even when the outcome has been reduced in value (Colwill & Rescorla, 1990). The procedure is sketched in Figure 1. Originally the animals were trained to nosepoke (Rc) in the

Discrimination	Target Train	Target Train	Deval	Test
L: Rc—O1	R1—O1	R1—O3	O1+	L:
				R1, R2
N: Rc—O2	R2—O2	R2—O3	O2+	N:

Figure 1. Experimental design for Experiment 1. Two stimuli, light (L) and noise (N), signaled whether a common nosepoke response (Rc) would produce one or the other outcome (O). Then leverpress and chainpull (counterbalanced as R1 and R2) first earned the differential outcomes, O1 and O2, and then the common O3. Each of the differential outcomes was paired with LiCL (+), and then L and N were separately presented with R1 and R2 to assess the state of the R-O associations.

presence of both a light and a noise, with different outcomes (O1 and O2) following responding during each stimulus. Then they were trained to make leverpresses and chainpulls (R1 and R2) for the same different outcomes. Eventually both O1 and O2 were paired with LiCl, a procedure well demonstrated to produce dramatic reduction in their value. However, because this operation can be expected to reduce the levels of both R1 and R2, the responses were also trained with a third outcome (O3, polycose), in an attempt to maintain responding. Previous results (e.g., Rescorla, 1991) have indicated that this training with a common third outcome will result in continued performance without adversely affecting the original R-O1 and R-O2 associations. Finally, the animals were tested for the ability of L and N to augment R1 and R2 selectively. Notice that successful differential transfer would indicate successful meditation by a devalued 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 × 20.3 × 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 deep 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 nosepoke. Access to these manipulanda could be blocked by covering the lever with a metal shield, retracting the chain through a hole in the ceiling, and covering the nosepoke opening with a jeweled lens. 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. Mounted on the inside wall of this shell were speakers that permitted the presentation of a white noise and an 1800-Hz tone, each measuring approximately 76 dB re 20 μN/m² against a background level of 62 dB. Also mounted on that wall was a 6-W bulb that could be illuminated to provide a light stimulus during the otherwise dark session. 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 or a 15% solution of nondairy creamer (Acme brand). Also attached to that food magazine was a dispenser containing 45-mg pellets (P. J. Noyes, Formula A).

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

Procedure

Initial training. On the 1st day, the animals received two 20-min magazine training sessions, the first containing 20 pellets, and the second, 20 sucrose presentations, delivered on a variable time (VT) 1-min schedule. Over the next 3 days, all animals were trained to make the three responses. Each training session allowed responding to earn 25 deliveries of the appropriate outcome on a continu-

ous reinforcement schedule. Half the animals received one session in which leverpressing led to pellets and one in which chainpulling led to sucrose; for the other half of the animals, the contingencies were reversed. Then all animals received two sessions in which the nosepoke response was reinforced. The outcome was pellets in the first session and sucrose in the second. Throughout this initial training, individual shaping was used if necessary for a particular R-O combination.

Discrimination training. On each of the next 2 days, the animals received two 20-min sessions, during which the nosepoke was rewarded on a variable interval (VI) 1-min schedule. One session of each day had pellets as the outcome and the other had sucrose. On each of the next 12 days, the animals received discrimination training sessions with nosepoke. Each session contained sixteen 30-sec trials with the light and sixteen 30-sec trials with the noise. During these stimuli, nose poking resulted in outcomes according to a VI 30-sec schedule. In a manner counterbalanced with regard to previous treatments, light signaled the availability of pellets and noise signaled the availability of sucrose for half the animals; the reinforcement contingencies were reversed for the other half of the animals. The intertrial intervals (ITIs) were variable around means of 15 sec, 30 sec, and 60 sec for the first 3 days of training, respectively. Thereafter, the mean ITI was 90 sec.

Target response training. On each of the next 5 days, the animals received VI 1-min training with both leverpressing and chainpulling. Each of two 20-min sessions, spaced about an hour apart, contained one manipulandum, and the outcomes were those used in initial response training. On each of the next 5 days, all animals continued to receive two 20-min VI sessions, one with lever and one with chain. The outcome earned in each of these sessions was 15% polycose. On the next 2 days, the animals were returned to the differential training with pellets in one session and sucrose in the other.

Discrimination retraining. On each of the next 3 days, the animals received additional discrimination training with L and N and the nosepoke response, exactly as in the original training.

Devaluation. Over the next 12 days, the animals received a treatment consisting of four 3-day cycles intended to devalue pellets and sucrose but leave polycose valuable. All manipulanda were removed from the chambers and the animals received pellets, sucrose, and polycose each on 1 day of each cycle. For the first three cycles, each outcome was delivered 20 times on a VT 1-min schedule. At the end of the pellet and sucrose sessions, the animal was removed from the chamber, given the LiCl injection, returned to its home cage, and fed 2 h later. The injection consisted of a .5% body weight .6 M solution of LiCl. On the days with polycose, the procedure was the same, but no injection was given. The order of presentation of the three outcomes was balanced across days. During the fourth cycle, the doors on the isolation chambers were left open to permit observation of the animals. On that day, an animal was removed and given its injection as soon as five reinforcer deliveries were left unconsumed.

Tests. On the next day, all animals received two test sessions, one with lever and one with chain. During each session, they received four 30-sec presentations each of N and L, spaced 30-sec apart. The data of interest are the differential amounts of transfer of N and L to these responses, based on whether or not they shared an outcome. On the next day, the animals received two 5-min consumption test sessions with all manipulanda removed. At the beginning of that session, two deliveries were given of either pellets or sucrose. The amount of outcome remaining after the 5-min session was recorded.

Results and Discussion

Initial discrimination training proceeded smoothly. By the final day of discrimination training, the mean response rates during the stimuli were 34.5 responses per minute;
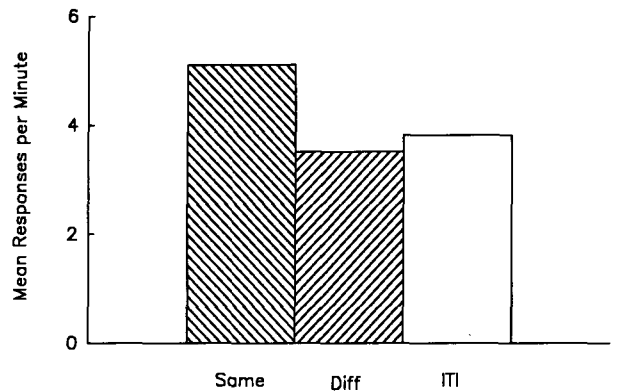
Figure 2. Responding during the test phase of Experiment 1. The mean rate of responding is shown in the absence of any stimulus (intertrial interval, ITI) and in the presence of stimuli that signaled the same or a different outcome.

those during the ITI were 3.5 responses per minute. On the last day of training of leverpress and chainpull with differential outcomes, the mean rate of responding was 12.4. Over the course of devaluation, both outcomes paired with LiCl came to be rejected, whereas the polycose continued to be consumed. On the final consumption test, the animals all left the sucrose; a total of three pellets were consumed across all of the animals.

The results of most interest, from the transfer test in which L and N were superimposed on leverpressing and chainpulling, are shown in Figure 2. That figure displays responding in the absence of any stimulus presentation (ITI) and in the presence of a stimulus that signaled either the same or the different outcome. It is clear that signaling the different outcome did little to elevate responding relative to the ITI. However, presentation of a same-outcome stimulus augmented responding relative to both the ITI [Wilcoxon $T(15) = 0, p < .01$] and the different-outcome stimulus [$T(15) = 21.5, p < .05$].

The pattern of results observed here is similar to that found in previous transfer experiments. Despite the fact that the outcomes had been devalued by pairing with LiCl, they continued to differentially mediate responding. That encourages the view that transfer does not require the continued value of the mediating outcome.

EXPERIMENT 2

Although Experiment 1 suggests that the continued high value of an outcome is not necessary to mediate transfer, it provides little evidence on whether that outcome value affects the magnitude of transfer. This experiment was an initial attempt to investigate that question. The design was like that of Experiment 1, except that only one of the outcomes was devalued. All animals were initially trained to nosepoke during L and N for the different O1 and O2 outcomes (pellets and sucrose). They then received differential training of R1 and R2 (leverpress and chainpull) with O1 and O2, followed by training of both responses

by the common O3. Eventually L and N were tested for transfer to R1 and R2. However, before that test, either O1 or O2 was paired with LiCl. These procedures permit a comparison of transfer by two stimuli to two responses under conditions in which the mediating outcome either remains valuable or has been devalued.

This design is very like that used by Colwill and Rescorla (1990), who reported successful transfer whether or not the outcome had been devalued. However, their design did not include the training with the common O3 outcomes. As a result, the overall levels of responding during the test were quite different for the responses that had and had not been paired with LiCl. That difference in baseline response rates made it impossible to compare directly the magnitudes of transfer obtained. The intention of using O3 in the present experiment was to produce more similar levels of responding and therefore allow that comparison.

Method

Subjects and Apparatus

The subjects were 16 rats like those in Experiment 1 and maintained in the same manner. One animal became ill during the experiment and was omitted from the analysis. The apparatus was that from Experiment 1.

Procedure

Initial response, discrimination, and transfer response training were all carried out as in Experiment 1. However, devaluation consisted of five 2-day cycles intended to devalue one outcome but not the other. On odd-numbered days, LiCl followed the outcome delivered; on even-numbered days, no LiCl was administered. For half the animals, pellets were paired with LiCl, whereas for the other half, sucrose was paired with LiCl. The identity of the outcome devalued was arranged to be orthogonal to the preceding treatments. The polycose O3 was not presented during this phase of the experiment.

The animals received a test day containing two sessions like those administered in Experiment 1. During the sessions, either lever or chain was present and they received four 30-sec presentations each of L and N, spaced 30 sec apart. On the next day, all animals received a 20-min choice test session during which both lever and chain were present but no outcomes were delivered. The intention of this test was to provide a direct choice between responses whose outcomes had been devalued and those whose values had not, so as to maximize the detection of any differences.

Results and Discussion

The various stages of training proceeded as in Experiment 1. At the end of discrimination training, the mean rates of responding were 23.6 and 2.0 responses per minute during the presence and absence of the stimulus, respectively. The mean response rate was 8.3 responses per minute at the end of transfer response training. Over the course of devaluation, the animals came to consume the outcomes differentially. In the final cycle of devaluation, the animals consumed all the nondevalued outcomes and all left the devalued outcomes.

The results of most interest, from the transfer test, are shown in Figure 3. That figure displays the mean rates for the responses whose outcomes had and had not been devalued, during the ITI, during a stimulus signaling the
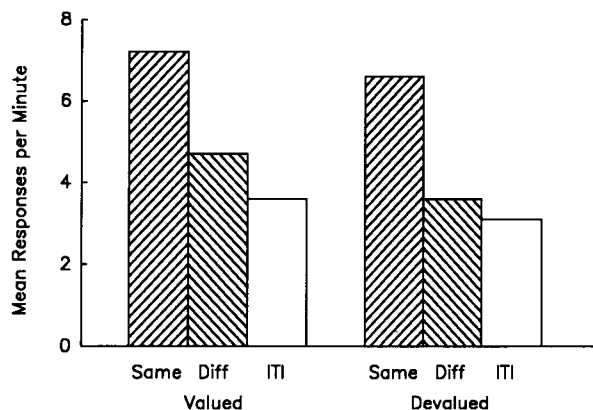

Figure 3. Responding during the test phase of Experiment 2. The right-hand bars show the rate for the response whose unique outcome had been devalued, in the absence of any stimulus (intertrial interval, ITI) and in the presence of stimuli that signaled the same or a different outcome. The left-hand bars show the rate for the response whose unique outcome had not been devalued for the same stimulus conditions.

same outcome, and during a stimulus signaling a different outcome. Both the overall levels of responding and the pattern of results are comparable for the devalued and nondevalued outcomes. In both cases, the same-outcome stimulus augmented responding relative to both that in the ITI and that during the different-outcome stimulus. Those differences proved reliable for both for the valued outcomes [$Ts(15) = 1$ and $7, ps < .01$] and for the devalued outcomes [$Ts(15) = 3$ and $5, ps < .01$]. Moreover, the magnitude of the superiority produced by the same-outcome stimulus was not different for the valued and devalued outcomes, when measured in terms of either difference scores [$T(15) = 68.5$] or ratios [$T(15) = 61.5$]. The numerically slightly greater overall responding on the nondevalued response during the test proved not statistically reliable, strengthening these comparisons. However, during the simple choice extinction test administered on the next day, the mean rate of responding was marginally higher for the nondevalued response [1.7 vs. 1.2 responses per minute; $T(15) = 28$ and $.05 < p < .10$].

These results are similar to those reported by Colwill and Rescorla (1990), with two exceptions. First, the present experiment was reasonably successful in matching the levels of responding and thus provides a stronger basis for concluding that devalued outcomes are just as effective as valued outcomes in mediating transfer. Second, unlike in the Colwill and Rescorla experiment, in the present experiment there was no evidence that a different-outcome stimulus depressed the level of responding, whether or not the outcome had been devalued. This difference seems likely to be attributable to differences in the details of the test procedures used. Colwill and Rescorla used a choice transfer test procedure so that both a same-outcome and different-outcome response were concurrently present during each stimulus. With such a test, augmentation of the same-outcome response may have a depressive effect on the different-outcome response. The single-response

test procedure used here would not have such response interactions.

The main finding of the present experiment, however, is that when the rates of responding are fairly comparable, transfer seems just as well mediated by a devalued outcome as by one that retains its value.

EXPERIMENT 3

This experiment was an attempt to evaluate further the role of outcome value in mediation of transfer, by using a somewhat more sophisticated design. The intention was to transfer various stimuli to responses under conditions in which each R and each S had been associated with two outcomes, one of which had been devalued. This allows more precise matching of the current values of the outcomes associated with the various responses and stimuli. But it also allows comparison of transfer mediated by valued and devalued outcomes from that common baseline.

Figure 4 illustrates the design. Initially, all animals learned to nosepoke during both L and N. Each stimulus signaled that responding would lead to one of two outcomes, either O1 and O2 or O3 and O4. The alternative outcomes were counterbalanced as pellets, sucrose, polycose, and 15% nondairy creamer. Then each animal learned separately to make two target responses, leverpress and chainpull. Again, each response earned one of two outcomes, in this case either O1 and O3 or O2 and O4. Next, two of the outcomes, O1 and O4, were paired with LiCl, whereas the other two were not. Finally, each stimulus was presented while either R1 or R2 was available. The feature of the design to note is that each S and each R has one devalued and one nondevalued outcome associated with it. Moreover, although each S shares one outcome with each R, in some cases the shared outcome has been devalued and in others it has not. For instance, the transfer of L to R1 is mediated by the devalued O1, but the transfer of L to R2 is mediated by the still-valued O2. Consequently, one can directly compare the success of transfer for responses and stimuli matched in overall value but differing in the value of the mediator.

Method

Subjects and Apparatus

The subjects were 16 naive rats like those in Experiment 1 and maintained in a similar manner. The apparatus was that from Ex-

Discrimination	Target Train	Devaluation	Test	
L: Rc−O1, O2	R1−O1, O3	O1+, O4+	L:	R1, R2
N: Rc−O3, O4	R2−O2, O4	O2−, O3−	N:	

Figure 4. Experimental design of Experiment 3. Two stimuli, light (L) and noise (N), signaled whether a common nosepoke response (Rc) would produce two outcomes, O1 and O2 or O3 and O4. Then two target responses, R1 and R2 (leverpress and chainpull), were both trained with two outcomes, either O1 and O3 or O2 and O4. Then one pair of outcomes, O1 and O4, was paired with LiCL (+). Finally, L and N were each presented separately with R1 and R2.

periment 1. In some sessions, .3 ml of 15% nondairy creamer was made available.

Procedure

Initial training. On the first 4 days, the animals received magazine training. All manipulanda were removed, and on each day they received 20 deliveries of an outcome on a VT 1-min schedule. The outcomes were pellets, sucrose, polycose, and creamer on Days 1–4, respectively. On the next 3 days, all animals were trained to leverpress, chainpull, and nosepoke. Each training session allowed responding to earn 25 deliveries of the appropriate outcome on a continuous reinforcement schedule. Half the animals received one session in which leverpressing led to pellets and one in which chainpulling led to sucrose; for the other half of the animals, the contingencies were reversed. Then all animals received two sessions in which the nosepoke response was reinforced. The outcome was pellets in the first session and sucrose in the second. Throughout this initial training, individual shaping was used if necessary for a particular R–O combination.

Target response training. Over the next 12 days, the animals were trained to leverpress and chainpull on a VI 1-min schedule. Each day contained two 20-min sessions, one with lever and one with chain. Half the sessions with each manipulandum had each of its two outcomes as the reward. For half the animals, lever led to pellets on six sessions and polycose on the other six; for the same animals, chain led to sucrose and creamer. For the other half of the animals, the outcome pairs were interchanged.

Discrimination training. Over the next 24 sessions, all animals received discrimination training with L and N. The nosepoke was present in all sessions and all contained sixteen 30-sec deliveries each of L and N. During each stimulus, reward was available on a VI 30-sec schedule. The ITIs on the first 4 days were 30, 30, 60, and 60 sec; thereafter, the ITI was 90 sec. For half the animals, L signaled pellets and creamer whereas N signaled sucrose and polycose; for the remaining half of the animals, the outcome pairs were interchanged. These treatments were balanced across animals so that they would be orthogonal to those of the preceding phase.

Target response retraining. On the next 2 days, the animals received VI training of the target responses in the same manner as before.

Devaluation. Over the next 16 days, the animals received four 4-day cycles of outcome delivery. During each session, the manipulanda were withdrawn and the animals received one outcome delivered on a VT 1-min schedule. Each outcome was given once during each cycle, and the sequence was scheduled so that LiCl was administered at the end of every other session. For half the animals, pellets and sucrose were devalued whereas polycose and creamer were presented without LiCl. For the other half of the animals, polycose and creamer were paired with LiCl. The selection of which outcomes to devalue was arranged to be orthogonal to the preceding treatments. In each of the first three cycles, the animals received 20 deliveries of the outcome. In the final cycle they received 10.

Test. On the next day, all animals received two test sessions, one with lever and one with chain. Each session contained four 30-sec presentations each of L and N, spaced 30 sec apart. The treatments were arranged so that for each manipulandum, each stimulus should augment responding; one of those stimuli should have its transfer mediated by a valued outcome and the other should have its transfer mediated by a devalued outcome.

Results and Discussion

Initial VI training of the transfer response proceeded smoothly, but there were some differences in the levels of responding for the various outcomes. On the final day of training with each outcome, the mean responses per
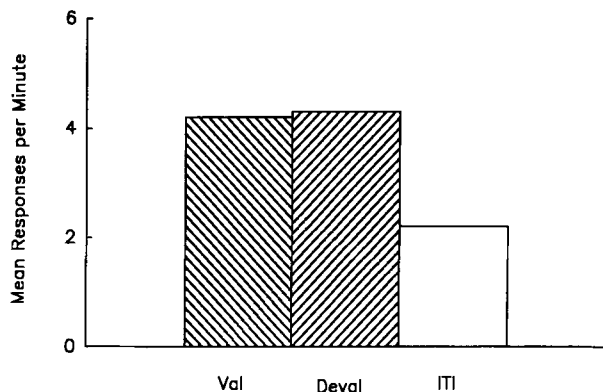
Figure 5. Results of the test session of Experiment 3. The rate of responding is shown during the intertrial interval (ITI) and during a stimulus that shared either a valued or devalued outcome with the target response.

minute were 9.0, 9.5, 7.1, and 4.1 for the pellets, sucrose, polycose, and creamer, respectively.

Discrimination training was without incident. In the final cycle of discrimination training, the mean number of responses per minute during the stimuli was 19, whereas that in the 30-sec period prior to a stimulus onset was 1.4.

Over the 16 days of devaluation training, the animals came to consume the outcomes differentially. In the final cycle, all animals consumed all the outcomes that had not been paired with LiCl, but all animals left in the wells some of the outcomes paired with LiCl.

The data of most interest, from the transfer test, are shown in Figure 5. That figure displays responding during three periods: in the absence of any stimulus, in the presence of a stimulus sharing a valued outcome with the response, and in the presence of a stimulus sharing a devalued outcome with the response. It is clear that both types of stimuli augmented responding, and to about the same extent. Responding was reliably greater during each stimulus than during the ITI [$Ts(16) < 13, ps < .01$]. However, there was no reliable difference in levels of responding during the stimuli [$T(14) = 41$].

This experiment provides a powerful comparison of transfer mediated by valued and devalued outcomes. Because of the experimental design, each response and each stimulus was associated with one outcome that was devalued and one that was not. Consequently, there was no opportunity for individual elements to differ in their overall value. Under those circumstances, the transfer seems to have been strong whether or not the mediating outcome was devalued.

One unfortunate aspect of the present experiment is that its design did not permit verification of outcome-specific mediation of transfer. It seems possible that the use of multiple outcomes, stimuli, and responses leads to generalization resulting in each stimulus's being able to augment an unusually broad range of responses. Although such a possibility cannot be ruled out, prior transfer ex-

periments (e.g., Rescorla 1991, 1992) have shown a remarkable preservation of outcome specificity with procedures similar in complexity to those used here. For instance, the present Experiment 2, although somewhat less complex in procedure, provided clear evidence of outcome specificity.

GENERAL DISCUSSION

The results of these experiments strongly suggest that the success of transfer of a discriminative stimulus to a new instrumental response does not depend on the current value of the outcome mediating that transfer. In Experiment 1, differential transfer survived devaluation of both a same and a different outcome. In Experiments 2 and 3, the levels of performance were matched to target responses in various ways, and little evidence was found to suggest that outcome value matters in controlling transfer. Apparently, as Colwill and Rescorla (1990) previously suggested, a discriminative stimulus will augment the rate of a target response on the basis of a shared outcome as long as the response is associated with *some* valued outcome.

It is, of course, quite difficult to be sure that the devaluation operations conducted here were complete in their reducing of the outcome value. Although the animals came to reject the outcomes, leaving them behind in a consumption test, it is possible that the outcomes retained some residual value. However, it is clear that the current operations did have a substantial effect on the current value of the outcomes. But that apparently did not change the amount of transfer.

These results are especially awkward for a view of transfer that sees the stimulus as making the outcome it shares with the response more accessible to the response. Such accessibility might occur because of an improved functioning of the R–O association or improved general availability of O. But in either case, a stimulus that acted in that manner would be expected to transfer much more successfully when the shared outcome was valuable.

These results are also out of line with a related intuitive description of transfer. Intuitively, discriminative training results in the stimulus signaling that a particular outcome is available upon the making of a response. Hence, if the animal has learned that R1 sometimes yields that outcome, then a stimulus that signals that this outcome is available might be expected to enhance R1. But that intuition envisions the animal as responding to obtain the mediating outcome. As a result it would erroneously anticipate substantial sensitivity to the current value of the outcome.

Mediation of transfer by a devalued outcome is much more compatible with an interpretation in terms of the stimulus activating an outcome representation that in turn activates the response. Under that account, the role of the mediating outcome is to serve as a step in a chain of events that activates the response. The stimulus activates a representation of the outcome, which in turn activates a rep-

resentation of associated responses. There is no inherent reason to anticipate that the value of the events in that chain has any particular significance.

Some have argued against a description of instrumental training in which the organism makes use of the R–O association in a "backward" direction (e.g., Mackintosh & Dickinson, 1979). However, the fact that the association is learned under conditions in which the response precedes the outcome need not imply that the association is routinely used in that same direction. In fact, it may be just in the case of instrumental learning that the organism needs to be able to activate potential outcomes and then use associations to recall what responses previously produced those outcomes. This may represent a difference in the manner of operation of Pavlovian S–O and instrumental R–O associations.

REFERENCES

ASRATYAN, E. A. (1974). Conditioned reflex theory and motivational behavior. *Acta Neurobiologiae Experimentalis*, **34**, 15-31.

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.

COLWILL, R. M., & RESCORLA, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 155-164.

COLWILL, R. M., & RESCORLA, R. A. (1990). Effect of reinforcer devaluation on discriminative control of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 40-47.

HOLLAND, P. C. (1983). "Occasion-setting" in Pavlovian feature positive discriminations. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Vol. IV. Discrimination processes* (pp. 183-206). Cambridge, MA: Ballinger.

MACKINTOSH, N. J., & DICKINSON, A. (1979). Instrumental (Type II) conditioning. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski* (pp. 143-169). Hillsdale, NJ: Erlbaum.

PETERSON, G. B., WHEELER, R. L., & ARMSTRONG, G. D. (1978). Expectancies as mediators in the differential-reward conditional discrimination performance of pigeons. *Animal Learning & Behavior*, **6**, 279-285.

RESCORLA, R. A. (1985). Inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299-326). Hillsdale, NJ: Erlbaum.

RESCORLA, R. A. (1990). Instrumental responses become associated with reinforcers that differ in one feature. *Animal Learning & Behavior*, **18**, 206-211.

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). Associations between an instrumental discriminative stimulus and multiple outcomes. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 95-104.

TRAPOLD, M. A., & OVERMIER, J. B. (1972). The second learning process in instrumental learning. In A. K. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 427-452). New York: Appleton-Century-Crofts.