

Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome

SEAN B. OSTLUND AND BERNARD W. BALLEINE
University of California, Los Angeles, California

We conducted three experiments to investigate the associative structure underlying the reinstatement of instrumental performance after extinction. In each experiment, rats were initially rewarded on two responses with different outcomes. At test, both responses were extinguished in order to assess the impact of a single noncontingent outcome delivery on response selection. Experiment 1 found evidence of outcome-selective reinstatement (i.e., more responses were performed on the lever that was trained with the reinstating outcome than on the other lever). Experiment 2 demonstrated that the outcome's capacity to reinstate performance was not affected by a reduction in its motivational value. Experiment 3 found evidence that the reinstating outcome selectively retrieved the response it signaled rather than the response it followed during training. Together, these findings are consistent with the view that instrumental reinstatement depends on the discriminative stimulus properties of the reinstating outcome.

It has been well established that, in instrumental conditioning, the training outcome plays a critical role in guiding action selection and initiation. For example, there is abundant evidence that posttraining outcome devaluation selectively suppresses the performance of an action trained with that outcome, relative to an action trained with a different outcome (e.g., Adams & Dickinson, 1981; Colwill & Rescorla, 1985). Moreover, studies of Pavlovian instrumental transfer have shown that the presentation of a conditioned stimulus will selectively facilitate the performance of an independently trained instrumental action with which it shares a common outcome (e.g., Colwill & Rescorla, 1988; Kruse et al., 1983). The outcome specificity of these effects discourages an explanation based on general motivational processes (e.g., satiety or arousal) and provides compelling evidence that action selection can be mediated by a detailed representation of the training outcome.

Several theories have been advanced to account for the involvement of outcome representations in action selection. One class of theory, derived from classic stimulus-response (S-R) theory, holds that the training outcome, or an expectation of that outcome, like other elements of the training situation, can become associated with a response if it reliably signals that the response will be rewarded (e.g., Capaldi, 1967; Trapold & Overmier, 1972). Thus, according to the simplest version of this account, it is the antecedent relationship between a previously earned outcome and the response next performed (i.e., an S_o -R association) that is encoded during training and used to direct

performance. In contrast, a second class of theory holds that it is the consequent relationship between an action and the outcome that it earns (i.e., a response-outcome [R-O] association) that is learned during training and that guides action selection (e.g., Asratyan, 1974; Bolles, 1972; Mackintosh & Dickinson, 1979). Because R-O theories preserve the teleological status of the outcome as a behavioral goal, it should not be surprising that most versions of this account propose that the capacity of the outcome to guide action selection depends on its current motivational value (but see Rescorla, 1994).

Despite their differences, both classes of associative theory predict that the mere presentation of an outcome will facilitate the performance of an associated response. Consistent with this prediction, it has long been known that noncontingently delivered outcomes can reinstate the performance of extinguished instrumental responding (e.g., Franks & Lattal, 1976; Reid, 1958; Rescorla & Skucy, 1969). Much remains unknown, however, about the specific function of the delivered outcome in instrumental reinstatement. It is possible, for example, that the outcome reinstates performance through one of several nonassociative mechanisms, including through the activation of nonspecific motivational or behavioral processes (e.g., hunger or arousal), or by disrupting any inhibition produced during extinction (i.e., through disinhibition). Although the basic reinstatement effect can be adequately explained by both associative and nonassociative accounts, the two accounts make very different predictions about the outcome specificity of reinstatement across dis-

S. B. Ostlund, sostlund@ucla.edu

tinct instrumental contingencies. Although both classes of associative theory predict that the noncontingent delivery of an outcome will selectively reinstate the action with which it was associated during training, nonassociative accounts of reinstatement tend to predict a general elevation in responding, regardless of the identity of the delivered outcome.

We conducted the present experiments in order to investigate these potential functions of the outcome in the reinstatement of instrumental performance. In each experiment, rats were trained to press two levers: Pressing one lever delivered sucrose solution and pressing the other lever delivered food pellets. After training, both leverpress actions were extinguished before a single outcome (either the pellet or sucrose) was delivered noncontingently and its effects on subsequent leverpressing were recorded. Experiment 1 assessed the outcome selectivity of instrumental reinstatement in order to evaluate the viability of associative accounts of this phenomenon. Experiments were then conducted to evaluate the two classes of associative theory: Experiment 2 assessed whether the effect of the outcome delivery on response selection depends on its current motivational value, and Experiment 3 investigated whether it is the antecedent or consequent association between the action and the outcome that mediates instrumental reinstatement.

EXPERIMENT 1

Evidence regarding the outcome selectivity of instrumental reinstatement has been mixed. There have been several reports of outcome-selective instrumental reinstatement after blocked training in which each response is rewarded with its unique outcome in a separate daily session (Colwill, 1994; Delamater, LoLordo, & Sosa, 2003; Leri & Stewart, 2001). However, only nonspecific reinstatement effects have been observed after training in which both responses are trained in the same daily session (Colwill, 1994; Delamater, 1997). We therefore used a blocked training procedure (see Table 1) in Experiment 1 to verify the outcome selectivity of instrumental reinstatement. Rats initially received free-operant training on two distinct R-O contingencies. Both responses (left and right leverpress) were then extinguished before we assessed the effect of a single noncontingent outcome delivery (either a food pellet or a drop of sucrose solution) on subsequent performance of the two leverpress actions in a choice extinction test.

Table 1
Design of Experiment 1

Training	Reinstatement Testing	
	Extinction	Reinstatement
R1-O1; R2-O2	R1-Ø; R2-Ø	O1: R1-Ø vs. R2-Ø

Note—During training, each of two responses (R1 and R2) earned a different outcome (O1 or O2). During the reinstatement test session, both responses were extinguished prior to the noncontingent delivery of a single outcome (e.g., O1), in order to assess its effects on subsequent performance. Ø, extinction (no outcome).

Method

Subjects and Apparatus

Sixteen adult female Long-Evans rats were individually housed in a temperature- and humidity-controlled vivarium that was illuminated on a 12:12-h light:dark cycle. Access to home chow was restricted in order to maintain subjects at 85% of their free-feeding body weight. Tap water was continuously available in the home cage throughout the experiment. The subjects were handled daily for 5 days prior to training.

The apparatus consisted of 16 identical operant chambers (Med Associates, East Fairfield, VT) that were housed within light- and sound-resistant shells. Centered at the base of one end wall was a food magazine, through which 45-mg food pellets (P. J. Noyes Co., Formula A/I) and 0.1-ml presentations of a 20% sucrose solution were dispensed. An infrared photo beam was positioned across the entrance to the food magazine in order to monitor head entries. Two retractable levers were located one on either side of the food magazine (i.e., left and right levers). Illumination was provided by a 24-V, 3-W light centered at the top of the opposite end wall. The chambers were controlled by microcomputers running the Med-PC program (Med Associates), located in an adjacent room.

Procedure

Training. The subjects received two sessions of magazine training on successive days, during each of which, 15 presentations of each outcome (pellets and sucrose) were delivered in random order according to a random time (RT) 1-min schedule. Two 30-min instrumental training sessions were then conducted on each of the next 11 days. The two responses (left and right leverpress) were trained in separate sessions. For half of the subjects, left leverpresses earned pellets and right leverpresses earned sucrose solution, whereas the remaining subjects were given the opposite R-O contingencies. The daily training sessions were separated by at least 30 min, and their order was alternated over days. Leverpressing was continuously reinforced (CRF) on the first 2 days of training. Over the next 9 days, the schedule of reinforcement was gradually shifted, with 3-day increments of random ratio 5 (RR5), RR10, and RR20.

Reinstatement testing. Both responses were available but had no scheduled consequence throughout the test session. After a 15-min period of extinction, a single outcome, either one pellet or 0.1 ml of sucrose solution, was scheduled for delivery as soon as 5 sec had elapsed without the performance of either response. This outcome delay procedure was used to avoid adventitious reinforcement by the noncontingent outcome delivery. The first magazine entry after the outcome delivery initiated a 3-min reinstatement period, during which responses on both levers were monitored in extinction. The identity of the delivered outcome was counterbalanced with respect to the instrumental contingencies.

Results and Discussion

One rat failed to acquire leverpressing and was excluded from the experiment. For the remaining rats, however, instrumental training proceeded without incident and resulted in robust rates of responding on both the lever rewarded with the outcome-to-be-delivered during reinstatement (33.2 presses/min) and on the other lever (34.4 presses/min). An ANOVA found no effect of response (reinst vs. other; $F < 1$).

Data from the extinction phase of the reinstatement test are presented in Figure 1, plotted as the mean number of leverpresses in successive 3-min bins. As can be seen, withholding reinforcement effectively reduced the performance of both responses and, again, to a similar degree. A response (reinst vs. other) \times bin (1-5) ANOVA conducted on these results yielded a main effect of bin

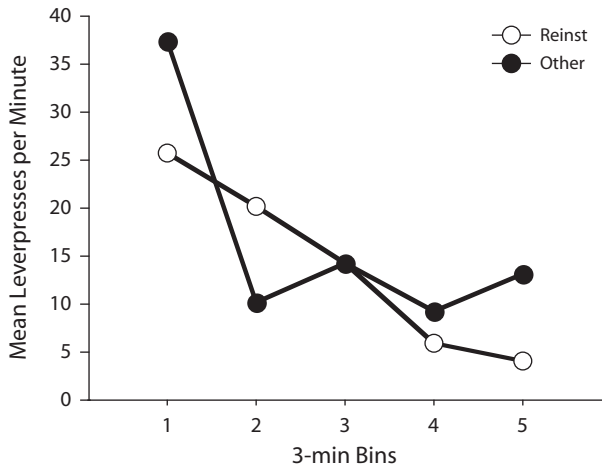


Figure 1. Experiment 1: Mean leverpresses per minute during the extinction phase of the reinstatement test, plotted separately for the response that had earned the reinstating outcome and the response that had earned the other outcome.

[$F(4,56) = 7.53, p < .0001$] but no effect of response, and no response \times bin interaction ($F < 1$).

The effect of the noncontingent outcome delivery on extinguished leverpressing was then assessed. The response rate data, expressed as the mean number of leverpresses performed per minute during the 3-min period that followed the outcome presentation, are presented in the left panel of Figure 2. The results clearly show that the outcome delivery selectively facilitated the response that had previously earned the reinstating outcome, relative to the other response. An ANOVA conducted on these data found a main effect of response (reinst vs. other) [$F(1,14) = 9.54, p < .01$].

Although analysis of the response rate data revealed the outcome selectivity of instrumental reinstatement, there was considerable variability across subjects in their

distribution of responses across levers during extinction, and in the overall magnitude of responding following the outcome delivery. Therefore, as a more direct measure of the effect of noncontingent outcome delivery on response selection, we also assessed the shift in choice performance across test phases. Specifically, data collected during the last 3 min of scheduled extinction (pre) and during the 3 min that followed the outcome presentation (post) were analyzed as the percentage of total responses made on the lever that, in training, delivered the reinstating outcome (i.e., $\text{reinst}/[\text{reinst} + \text{other}] \times 100$). These data, presented in the right panel of Figure 2, clearly demonstrate that subjects increased their selection of the response that had earned the reinstating outcome after it was delivered. This conclusion was supported by an ANOVA performed on the data, which found a main effect of test phase (pre vs. post) [$F(1,14) = 18.5, p < .001$].

These results confirm earlier reports of outcome-selective reinstatement (Colwill, 1994; Delamater et al., 2003; Leri & Stewart, 2001), a finding that implies that the noncontingent outcome has its influence over instrumental performance through an associative response-retrieval process. If the effect were to depend instead on the activation of a nonspecific motivational or behavioral process (e.g., a general increase in hunger or arousal), or some form of disinhibition, then the performance of both actions should have increased, and no specific change in choice performance should have been observed.

Nevertheless, because the levers were available throughout the test session, including the time of the outcome presentation, it remains possible that the outcome had its effect on performance, not by triggering response selection and initiation, but by adventitiously rewarding the last leverpress that was made during extinction. This account seems unlikely for two reasons. First, a delay was imposed between residual responding and the noncontingent outcome delivery in order to prevent an unintended pairing between these events. Second, post hoc analysis of perfor-

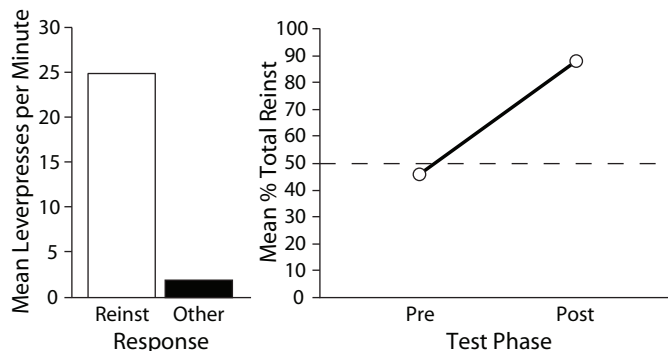


Figure 2. Experiment 1: The left panel shows the mean leverpresses per minute performed following the reinstating outcome delivery, plotted separately for the response that had earned that outcome and the response that had earned the other outcome. The right panel shows test performance during the last 3 min of scheduled extinction (pre) and during the 3-min reinstatement phase (post), plotted as the mean percentage of total responses made on the lever that had earned the reinstating outcome.

mance during the last minute of recorded extinction found no effect of response (reinst vs. other; $F < 1$), suggesting that the specific reinstatement observed in this experiment was not the result of an adventitious R–O pairing, but was mediated instead by an association acquired during instrumental training.

EXPERIMENT 2

The results of Experiment 1 provide evidence for the involvement of an outcome-mediated response-selection process in instrumental reinstatement, a conclusion that is clearly compatible with several major models of instrumental performance. These theories make different assumptions, however, about the role played by the mediating outcome representation in response selection. Generally, R–O theories place emphasis on the role of the outcome as a behavioral goal. The outcome's capacity to initiate performance is therefore often explicitly assumed to depend on its current motivational value (e.g., Asratyan, 1974; Bolles, 1972; Mackintosh & Dickinson, 1979). In line with this view, it has been well established that instrumental performance can be suppressed by devaluing the outcome after training (e.g., Adams & Dickinson, 1981; Colwill & Rescorla, 1985). However, if the associative process underlying the influence of outcome value over performance is also responsible for the reinstatement effect, then the outcome's capacity to reinstate performance should depend on its current value. Alternatively, the S_o –R account assumes that the outcome mediating reinstatement acts essentially as a discriminative stimulus (e.g., Capaldi, 1967; Franks & Lattal, 1976; Reid, 1958; Rescorla & Skucy, 1969). According to this view, it is the sensory properties of the outcome, and not its incentive properties, that become associated with the instrumental response. Thus, the outcome's capacity to activate an associated response should be impervious to posttraining manipulations of its motivational value.

In order to compare these two accounts, Experiment 2 assessed the sensitivity of instrumental reinstatement to posttraining outcome devaluation. As depicted in Table 2, hungry rats were trained as in Experiment 1, with one response earning food pellets and the other response sucrose solution. Reinstatement testing was then conducted in rats that were either hungry, such that both outcomes maintained a high motivational value, or thirsty (but not hungry), in which case the sucrose solution should have been valued more than dry food pellets. Whereas the R–O account predicts that this shift from hunger to thirst should diminish the impact of the food pellet, but not the sucrose solution, on response selection, the S_o –R account predicts that it should have no such effect.

Method

Subjects and Apparatus

Thirty-two adult female rats of the same stock, maintained in the same manner as described in Experiment 1, were used as subjects. The apparatus was the same as that described in Experiment 1.

Table 2
Design of Experiment 2

Training	Reinstatement Testing	
	Extinction	Reinstatement
H: R1–P; R2–S	H: R1–Ø; R2–Ø	P: R1–Ø vs. R2–Ø
H: R1–P; R2–S	H: R1–Ø; R2–Ø	S: R1–Ø vs. R2–Ø
H: R1–P; R2–S	T: R1–Ø; R2–Ø	P: R1–Ø vs. R2–Ø
H: R1–P; R2–S	T: R1–Ø; R2–Ø	S: R1–Ø vs. R2–Ø

Note—During training, each of two responses (R1 and R2) earned a different outcome: either food pellets (P) or sucrose solution (S). Subjects underwent reinstatement testing either hungry (H) or thirsty (T). The reinstating outcome was either a pellet or sucrose solution. Ø, extinction (no outcome).

Procedure

Training. Magazine and instrumental training were conducted using the same procedures described in Experiment 1.

Motivational shift. During the 48 h interposed between the last day of instrumental training and reinstatement testing, the rats were either maintained on the same food-deprivation schedule used during training (hungry), or were provided with ad-lib access to home chow, but were restricted to 1 h of access to tap water per day (thirsty).

Reinstatement testing. Reinstatement testing was conducted using the same procedures described in Experiment 1, except that the extinction phase of the session was extended to 20 min. A food pellet served as the reinstating outcome for half of the rats in each test state, whereas the remaining rats received sucrose solution. Both test state (hungry vs. thirsty) and reinstating outcome (pellet vs. sucrose) were counterbalanced with respect to training contingency.

Results and Discussion

Given the relevance of outcome identity to the present experiment, the results were analyzed according to whether a given response earned pellets or sucrose. One subject failed to acquire leverpressing and was excluded from the experiment. By the end of training, the mean rate of responding for pellets (42.5 presses/min) did not differ significantly from that for sucrose (43.4 presses/min; $F < 1$). Moreover, group assignments (hungry, $n = 16$; thirsty, $n = 15$) were made on the basis of instrumental performance in order to control for response bias, such that we observed neither an effect of group nor a group \times response interaction ($F_s < 1$).

In contrast, instrumental performance during the extinction phase of the test session, which followed the motivational shift, did critically depend on group assignment. Figure 3 clearly shows that, whereas subjects tested when hungry responded at similar levels across levers, subjects tested when thirsty performed fewer responses on the lever that had earned pellets than on the lever that had earned sucrose. Furthermore, the extinction procedure resulted in considerable suppression of responding for all groups over the course of the 20-min session. These conclusions were confirmed by statistical analysis of the extinction data, collected in 4-min bins, using a response (pellet and sucrose) \times state (hungry vs. thirsty) \times bin (1–5) ANOVA. This test found a significant main effect of state [$F(1,29) = 6.57, p < .05$] and bin [$F(4,116) = 12.97, p < .0001$], but not response [$F(1,29) = 2.75, p > .05$]. The ANOVA also revealed significant response \times

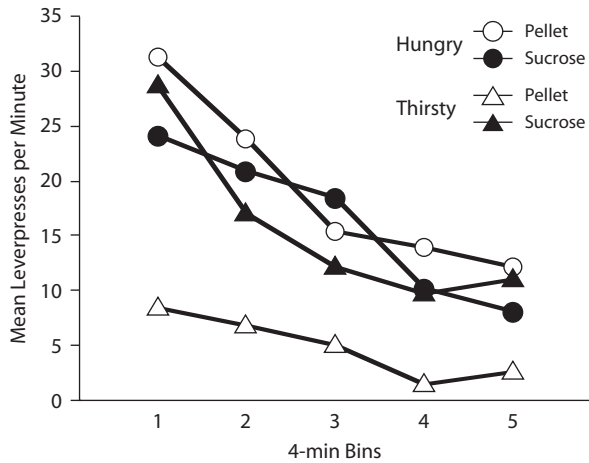


Figure 3. Experiment 2: Mean leverpresses per minute during the extinction phase of the reinstatement test, plotted separately for the response that had earned pellets and the response that had earned the sucrose.

state [$F(1,29) = 8.31, p < .01$] and response \times state \times bin [$F(4,116) = 2.55, p < .05$] interactions. Further analysis of the response \times state interaction indicated that, whereas those subjects tested when hungry displayed no response preference ($F < 1$), subjects tested when thirsty responded more for sucrose than for pellets [$F(1,14) = 14.01, p < .01$].

The response rate data from the reinstatement phase of the test session are presented in the left panel of Figure 4. In general, subjects tested when hungry were more likely to perform the response that had been trained with the reinstating outcome, regardless of the identity of that outcome (pellet, $n = 8$; sucrose, $n = 8$). Subjects tested thirsty, however, displayed a more complex pattern of responding. Although the group that received sucrose ($n = 7$) tended to perform more responses for sucrose than for pellets, those that received a food pellet ($n = 8$) displayed

no clear response preference. Given their preference for sucrose over pellets during extinction, however, it seems likely that the thirsty group's performance following the outcome delivery was heavily influenced by an underlying response bias. As a consequence, further analysis placed emphasis on the shift in choice performance across test phases (see below). Nevertheless, statistical analysis of the response rate data using a response (reinst vs. other) \times state (hungry vs. thirsty) \times outcome (pellet vs. sucrose) ANOVA revealed a significant main effect of response [$F(1,27) = 9.82, p < .01$] and state [$F(1,27) = 6.26, p < .05$], but not outcome [$F(1,27) = 1.18, p > .05$]. Although no interaction reached significance, the ANOVA did result in marginal response \times state [$F(1,27) = 3.09, p = .09$], state \times outcome [$F(1,27) = 3.25, p = .08$], and response \times state \times outcome interactions [$F(1,27) = 3.26, p = .08$].

Although this trend might be taken as evidence that the motivational shift from hunger to thirst left the food pellet ineffective in reinstating its associated response, analysis of the choice performance data, presented in the right panel of Figure 4, appears to demand a rather different interpretation. As is clear from this figure, all groups increased their performance of the response that had earned the reinstating outcome relative to their baseline choice performance during extinction. A test phase (pre vs. post) \times state (hungry vs. thirsty) \times outcome (pellet vs. sucrose) ANOVA found a main effect of phase [$F(1,27) = 8.83, p < .01$] and outcome [$F(1,27) = 8.91, p < .01$], but not state ($F < 1$). Although the main effect of outcome indicates that subjects tested with food pellets were, in general, less likely to choose the response that had earned that outcome, the significant state \times outcome interaction [$F(1,27) = 20.66, p < .0001$] revealed that this effect was largely carried by the thirsty group. Simple effects analysis found a significant effect of outcome in the thirsty condition [$F(1,13) = 38.29, p < .0001$] but not in the hungry condition ($F < 1$). More importantly, however, no other effect or interaction in the main ANOVA reached signifi-

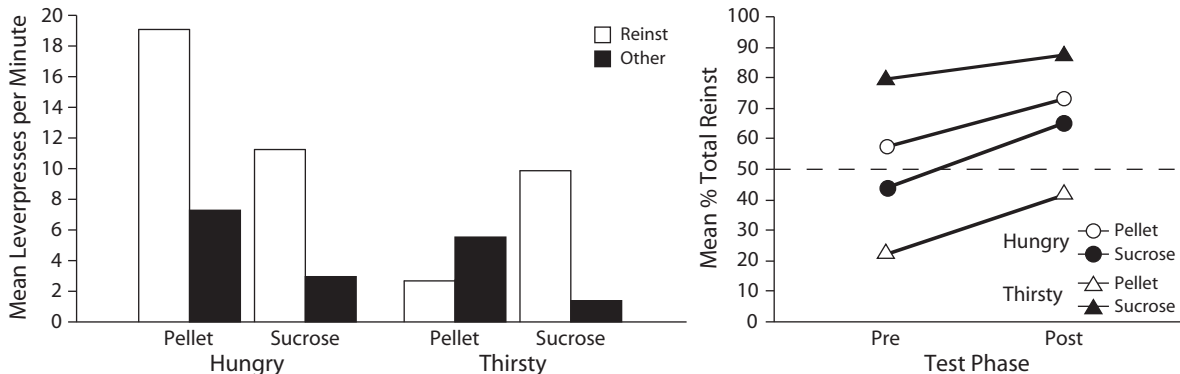


Figure 4. Experiment 2: The left panel shows the mean leverpresses per minute performed following the reinstating outcome delivery, plotted separately for the response that had earned that outcome and the response that had earned the other outcome. The right panel shows test performance during the last 3 min of scheduled extinction (pre) and during the 3-min reinstatement phase (post), plotted as the mean percentage of total responses made on the lever that had earned the reinstating outcome.

cance ($F_s < 1$), suggesting that the shift in choice performance following the outcome delivery did not depend on test state, outcome identity, or an interaction between these factors.

The present findings provide a behavioral dissociation between the effects of a shift in outcome value on instrumental performance and the response-reinstating properties of that outcome. If these phenomena were to depend on a common mechanism, as is predicted by various R–O theories (e.g., Asratyan, 1974; Bolles, 1972; Mackintosh & Dickinson, 1979), then outcome devaluation should have abolished the reinstatement effect. The observed results are therefore more compatible with the S_o –R account (e.g., Capaldi, 1967; Franks & Lattal, 1976; Reid, 1958; Rescorla & Skucy, 1969), which attributes the reinstatement effect to the discriminative stimulus properties of the outcome.

Although R–O theories typically assume that the outcome's capacity to retrieve an associated response is determined by that outcome's motivational value, Rescorla (1994) has noted that this is not a necessary assumption of R–O theory and has briefly described an alternative view, in which these two properties of the outcome are considered independent of one another. This view, although originally applied to the Pavlovian instrumental transfer effect, can be extended to provide an adequate explanation of the present results. Therefore, although the results of Experiment 2 are suggestive, they do not provide a critical test of the associative structure underlying instrumental reinstatement.

EXPERIMENT 3

Experiment 3 was designed to provide a more decisive comparison between the S_o –R and R–O accounts of reinstatement. Both theories posit that reinstatement is mediated by a previously acquired association between the instrumental response and the reinstating outcome. However, these theories differ fundamentally in terms of the relationship thought to be encoded in that association. Whereas the S_o –R account holds that the outcome is encoded as an antecedent to the response, the R–O account holds that it is encoded as a consequent event.

Previous failures to observe outcome-selective reinstatement (Colwill, 1994; Delamater, 1997) may provide insight into the directional content of the underlying association. For instance, Colwill (1994) observed only nonspecific reinstatement after concurrent training on two different R–O contingencies, such that subjects were free to alternate between the two responses. Importantly, unlike the blocked training procedure used in Experiments 1 and 2, concurrent training of this kind prevents subjects from experiencing stable O–R relationships during acquisition. Similarly, Delamater (1997) found nonspecific instrumental reinstatement after training rats on a biconditional discrimination in which R–O, but not O–R, relationships were stable. Thus, in line with the S_o –R account, such findings suggest that reliable antecedent, O–R relationships are critical for selective reinstatement.

In order to more directly assess this account, we adapted a training procedure used previously to investigate the directional content of the association underlying outcome devaluation and transfer (Rescorla, 1992; Rescorla & Corwill, 1989). As depicted in Table 3, rats were trained using a discrete-trial procedure in which each trial was initiated by the noncontingent delivery of a single outcome (either a pellet or a drop of sucrose). After that outcome was consumed, one of the two levers was inserted into the chamber. Pressing this lever delivered either the same outcome, for the congruent training group, or the other outcome, for the incongruent training group. As with the blocked training procedure used in Experiments 1 and 2, congruent training ensured that each outcome held an antecedent and consequent relationship with the same response, and held neither relationship with the other response. Hence, although it was predicted that the congruent group would display clear and selective reinstatement performance, the results from this group would not be informative as to the directional nature of the association mediating this effect. For the incongruent training group, however, each outcome held an antecedent relationship with one response and a consequent relationship with the other response, making it possible to contrast the relative contribution of S_o –R and R–O associations to reinstatement performance. If, as suggested by the results of Experiment 2, selective

Table 3
Design of Experiment 3

Training	Reinstatement Testing		Devaluation Testing
	Extinction	Reinstatement	
O1–R1–O1	R1–Ø; R2–Ø	O1: R1–Ø vs. R2–Ø	O2 (sate): R1–Ø vs. R2–Ø
O2–R2–O2			
O2–R1–O1	R1–Ø; R2–Ø	O1: R1–Ø vs. R2–Ø	O2 (sate): R1–Ø vs. R2–Ø
O1–R2–O2			

Note—During training, each of two responses (R1 and R2) earned a different outcome (O1 or O2). Unlike in earlier experiments, however, each response was also signaled by a noncontingent presentation of either the outcome that it earned (congruent training) or the outcome that was earned by the other response (incongruent training). Reinstatement testing was followed by outcome devaluation testing, in which subjects were sated on one outcome (e.g., O2), in order to assess the effects of this treatment on subsequent performance. Ø, extinction (no outcome).

reinstatement of instrumental responding is mediated by an S_0 -R association, then reinstatement should be greater for the response *signaled* by the delivered outcome. Alternatively, if this effect is mediated by an R-O association, then reinstatement should be greater for the response that had previously *earned* the delivered outcome.

Although little is known about the nature of the associative structure underlying reinstatement, there is good evidence that the sensitivity of instrumental performance to outcome devaluation is mediated predominantly by an R-O association (for a review, see Balleine & Dickinson, 1998; Colwill & Rescorla, 1986). Rats in the present study were therefore also administered an outcome devaluation test to verify these earlier findings and further assess the potential dissociability of the processes underlying outcome devaluation and reinstatement. If outcome devaluation performance does, in fact, depend on an R-O association, then both groups should have made fewer presses on the lever that had previously *earned* the devalued outcome than on the other lever, regardless of which outcome *signaled* these levers during training.

Method

Subjects and Apparatus

Sixteen adult female rats of the same stock, maintained in the same manner as described in Experiment 1, were used as subjects. The apparatus was the same as that described in Experiment 1.

Procedure

Training. Magazine training was conducted in the same manner as was described in Experiment 1. Two daily instrumental training sessions were conducted on each of the next 11 days. Once again, each response was rewarded with a different outcome. For half of the subjects, left leverpresses earned pellets and right leverpresses earned sucrose solution, whereas the remaining subjects received the opposite instrumental contingencies. Reinforcement was scheduled in the same manner as in Experiment 1, with 2 days of CRF and 3 days each at RR5, RR10, and RR20. Training was conducted using an adapted discrete-trial procedure in which each trial was initiated by the noncontingent delivery of one of the two outcomes. Four seconds after that outcome was retrieved, as measured by magazine entry, the appropriate lever was inserted into the chamber. The lever remained extended until the response requirement (determined by the current reinforcement schedule) had been met. Each response was reliably signaled by a different outcome. For the congruent group, the antecedent and consequent outcomes for each response were the same (e.g., pellet: left leverpress → pellet), whereas for the incongruent group, these outcomes were different (e.g., pellet: left leverpress → sucrose). Group and instrumental contingency assignments were counterbalanced. As in Experiment 1, the two responses were trained in separate daily sessions. Within each session, individual trials were separated by a 90-sec variable intertrial interval (range: 60–120 sec), during which the lever was retracted. Each session ended after 15 outcomes had been earned. Session order was alternated over days.

Reinstatement testing. Two sessions of reinstatement testing were given, using procedures similar to those described in Experiment 1. In the present experiment, however, the extinction period of the test session was extended to 20 min (as in Experiment 2) and the postdelivery period was extended to 6 min. Each test involved the delivery of a different reinstating outcome. Test order (i.e., pellet/sucrose vs. sucrose/pellet) was counterbalanced with group assignment and instrumental contingency. A session of retraining (RR20) on each response was given on the day between reinstatement tests.

Devaluation testing. A session of retraining (RR20) on each response was given on the day between reinstatement and devaluation

testing. The rats were given two sessions of devaluation testing. In each test, rats were initially given 1 h of unrestricted access to either sucrose solution (30 ml in a graduated glass drinking tube) or pellets (30 g in a small glass bowl) in their home cage in order to devalue that outcome through sensory-specific satiety. Immediately after this treatment, the rats were returned to the experimental chambers for a 10-min choice extinction test, during which both response levers were inserted into the box but no outcomes were delivered. Each test involved the devaluation of a different outcome. The test order used during devaluation testing was reversed, relative to reinstatement testing. Rats were given a session of retraining (RR20) on each response on the day between devaluation tests.

Results and Discussion

One rat in the congruent training group failed to acquire leverpressing and was excluded from the experiment. Training was otherwise uneventful, and the two groups acquired similar levels of instrumental performance; the average total session length (i.e., the time that elapsed before 15 outcomes were earned, including intertrial intervals) on the last day of training for the congruent group (32.9 min) did not significantly differ from that for the incongruent group (34.0 min; $F < 1$).

It is worth noting that the discrete-trial procedure used in this experiment resulted in lower rates of responding during the reinstatement test (see below) than was observed in Experiments 1 and 2. There were several procedural differences across experiments that may account for this effect (e.g., the total number of reinforced responses, the delivery of noncontingent outcomes during training, and the imposition of time-out periods between trials). Importantly, however, the within-subjects design used here allowed our data analysis to target the associative component of reinstatement without being hindered by fluctuations in overall response rates across subjects or experiments.

The results of the extinction phase of the reinstatement test are presented in Figure 5, plotted according to whether the response had previously earned the rein-

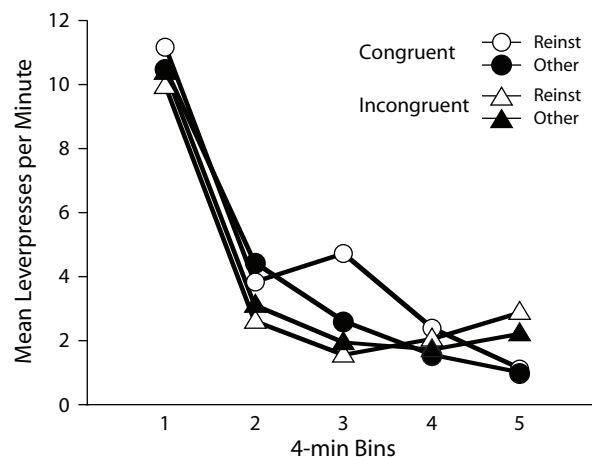


Figure 5. Experiment 3: Mean leverpresses per minute during the extinction phase of the reinstatement test, plotted separately for the response that had earned the reinstating outcome and the response that had earned the other.

stating outcome (reinst) or the other outcome (other). As can be seen, both groups suppressed their performance as the extinction phase progressed, and to a similar degree on both responses. A response (reinst vs. other) \times group (congruent vs. other) \times bin (1–5) ANOVA revealed a significant effect of bin [$F(4,52) = 39.1, p < .0001$]. No other main effect or interaction reached significance [highest $F(4,52) = 1.34, p > .05$].

The response rate data from the reinstatement phase of the test, presented in the left panel of Figure 6, show that the two groups displayed strikingly different patterns of responding following the noncontingent outcome delivery. As predicted, the congruent group performed more responses on the lever that had previously earned the reinstating outcome during training than on the other lever. In contrast, the incongruent group displayed the opposite pattern of results, performing fewer responses on the lever that had earned the reinstating outcome than on the other lever, which, for the incongruent group, had been signaled by that outcome during training. A response (reinst vs. other) \times group (congruent vs. incongruent) ANOVA found no main effect of either response ($F < 1$) or group [$F(1,13) = 1.90, p > .05$], but did reveal a significant response \times group interaction [$F(1,13) = 7.76, p < .05$]. Further analysis revealed a marginal effect of response for both the congruent [$F(1,6) = 4.61, p < .08$] and incongruent group [$F(1,7) = 3.55, p = .10$].

The right panel of Figure 6 presents the test data as the mean percentage of total responses performed on the lever that had previously earned the reinstating outcome during the last 6 min of extinction (pre) and during the 6 min that followed the outcome delivery (post). Although choice of the response that had earned the reinstating outcome increased in the congruent group following its delivery, the incongruent group decreased their choice of this response. A group (congruent vs. incongruent) \times phase (pre vs. post) ANOVA revealed a main effect of group [$F(1,13) = 13.56, p < .01$] but not phase ($F < 1$). More importantly, however, the analysis found a significant group \times phase

interaction [$F(1,13) = 8.97, p < .05$], indicating that the shift in choice performance differed across groups. Further analysis revealed a simple effect of group on choice performance during the postdelivery phase [$F(1,13) = 17.43, p < .01$], but not during the predelivery phase ($F < 1$). These results suggest that the influence of the reinstating outcome on choice performance depends on the antecedent relationship that existed between that outcome and the response during training; both groups selected the response that had previously been signaled by the reinstating outcome, regardless of whether that response had earned that outcome or a different outcome.

The results of devaluation testing are presented in Figure 7 as the mean number of responses performed per min, plotted separately for the response that had earned the now devalued outcome (deval) and for the other response (other). Although both groups selectively suppressed their performance of the response that had earned the devalued outcome, relative to the other response, congruent training seems to have led to a more pronounced devaluation effect. A response (deval vs. other) \times group (congruent vs. incongruent) ANOVA found a significant main effect of response [$F(1,13) = 22.91, p < .001$] but not group [$F(1,13) = 1.47, p > .05$]. The analysis also revealed a marginal response \times group interaction [$F(1,13) = 4.33, p = .06$]. Analysis of this interaction, however, revealed that both the congruent group [$F(1,6) = 14.33, p < .01$] and the incongruent group [$F(1,7) = 7.33, p < .05$] showed a significant effect of response, indicating that both groups selectively reduced performance of the response that had previously earned the devalued outcome. This pattern of results is consistent with the view that the influence of anticipated outcome value over performance is primarily mediated by an R–O association, as opposed to an S_o–R association. It should be noted that previous support for this claim has come from the finding that even after concurrent training on two distinct instrumental contingencies, such that each response is just as likely to be signaled by the outcome that it earns as by the

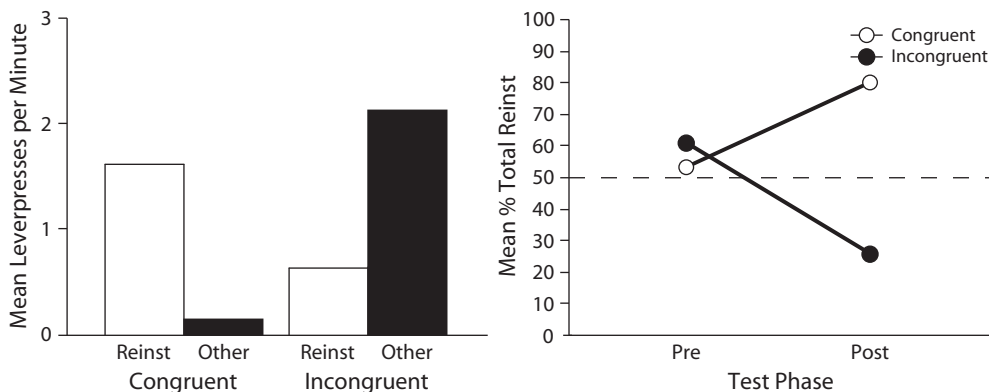


Figure 6. Experiment 3: The left panel shows the mean leverpresses per minute performed following the reinstating outcome delivery, plotted separately for the response that had earned that outcome and the response that had earned the other outcome. The right panel shows test performance during the last 6 min of scheduled extinction (pre) and during the 6-min reinstatement phase (post), plotted as the mean percentage of total responses made on the lever that had earned the reinstating outcome.

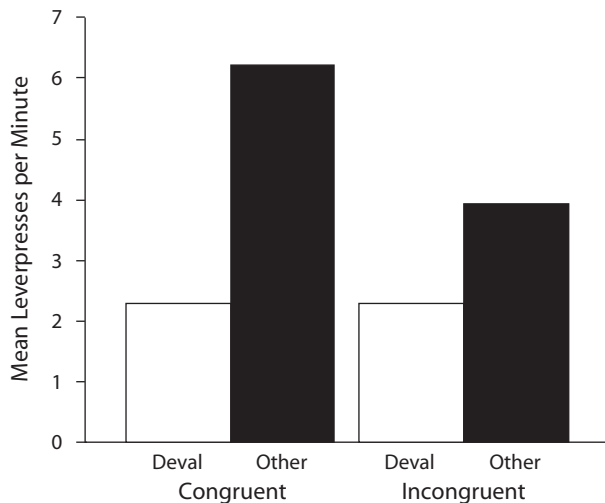


Figure 7. Experiment 3: Mean leverpresses per minute performed during the outcome devaluation test, plotted separately for the response that had earned the devalued outcome and the response that had earned the other outcome.

other outcome, performance remains sensitive to a shift in the *earned* outcome's value (Colwill & Rescorla, 1985, Experiment 2). Evidence of this kind, however, does not preclude the possibility that the O–R relationship experienced during training plays some, albeit minor, role in devaluation performance. Indeed, tentative support for this conclusion can be found in the present results. Although both groups showed a clear outcome devaluation effect, this effect was at least numerically smaller in the incongruent training group, for which the devalued outcome was earned by one response but signaled the other response during training. This finding mirrors an earlier observation that the size of the devaluation effect is smaller after concurrent training than after blocked training with stable R–O and O–R relationships (Colwill & Rescorla, 1985, Experiments 1 and 2). Although such findings are suggestive, further research will be needed to directly examine the influence of the O–R relationship experienced during training on outcome devaluation performance.

Taken together, these results indicate that both S_o –R and R–O associations are encoded during instrumental training but underlie distinct response-selection processes; whereas a noncontingently delivered outcome appears to reinstate performance through an S_o –R association, the sensitivity of instrumental performance to a reduction in *expected* outcome value depends primarily on an R–O association.

GENERAL DISCUSSION

The present series of experiments sought to characterize the associative structure underlying the reinstatement of previously extinguished instrumental performance. In Experiment 1, we observed that the noncontingent delivery of a single outcome selectively reinstated the response with which it had been paired during training, relative to a response that had been paired with a different outcome, indicating that reinstatement involves a response-selection

process mediated by a detailed outcome representation. The results of Experiment 2 suggest that an outcome's capacity to reinstate performance does not depend on its current value, consistent with the view that reinstatement is mediated by the discriminative stimulus properties of the outcome (i.e., through an S_o –R association). Experiment 3 provided further evidence for this conclusion; the delivery of a noncontingent outcome was found to selectively reinstate the response it had signaled during training, regardless of whether or not it had also been earned by that response.

So far, we have limited our discussion to the most straightforward version of the S_o –R view of reinstatement, according to which the outcome itself becomes encoded as a discriminative stimulus during training (e.g., Capaldi, 1967; Reid, 1958; Rescorla & Skucy, 1969). Unlike other elements of the training context, however, the outcome is absent during extinction and therefore retains its excitatory influence over performance. Thus, according to this account, reinstatement is mediated by direct association between the noncontingent outcome and the response it signaled during training. Alternatively, if we assume that subjects are sensitive to the incidental Pavlovian relationship between outcomes during training, it is possible to imagine a two-process account (cf. Trapold & Overmier, 1972) in which the outcome reinstates performance through an indirect connection—that is, the noncontingent outcome may evoke an expectation of the outcome it signaled during training, which in turn activates the response rewarded in the presence of that outcome expectation. Notice that this account assumes that reinstatement depends on a facilitatory Pavlovian instrumental interaction akin to that observed in the transfer effect. Indeed, reinstatement and transfer share many common features. Both effects are triggered by the noncontingent presentation of some event (either the outcome itself or a stimulus paired with that outcome), typically after some period of extinction. Both reinstatement (e.g., Experiments 1–3; Colwill, 1994; Delamater et al., 2003; Leri & Stewart, 2001) and transfer (e.g., Colwill & Rescorla, 1988; Kruse et al., 1983) are mediated by a detailed outcome representation. Finally, both effects appear to be insensitive to manipulations of the mediating outcome's value. The results of Experiment 2, for instance, show that whereas a shift in motivational state from hunger to thirst considerably reduced the baseline rate of responding for pellets relative to sucrose during the extinction period, it had no detectable effect on the pellet's capacity to reinstate extinguished performance. Similarly, there have been several demonstrations that lithium chloride–induced outcome devaluation has no impact on the outcome-specific transfer effect (Colwill & Rescorla, 1990; Holland, 2004; Rescorla, 1994).

Although the similarities between reinstatement and transfer might be taken as support for the two-process account, there is evidence that, unlike reinstatement, transfer is primarily mediated by an R–O association (e.g., Colwill, 1994; Rescorla, 1992; Rescorla & Colwill, 1989). For instance, rats trained concurrently with two stable R–O, but not O–R, relationships display outcome-selective transfer (Colwill, 1994) and outcome-devaluation (Colwill

& Rescorla, 1985) effects, but do not show selective reinstatement (Colwill, 1994). Thus, when considered together with the results reported here, these findings argue against accounts that necessarily assume that reinstatement and transfer are determined by the same associative process.

Finally, it is worth noting that, although instrumental reinstatement appears to depend primarily on an S_0 -R process, this does not appear to be the case for the reinstatement of Pavlovian conditioned responding. In order to contrast the outcome selectivity of instrumental and Pavlovian reinstatement, Delamater (1997) trained rats using a biconditional discrimination procedure in which each of two discriminative stimuli signaled that a different outcome could be earned on a common response manipulandum. In this study, however, trials were slightly more likely to alternate than repeat, so that the outcome earned on any given trial was not a reliable signal of either trial type. At test, noncontingent presentations of one of the training outcomes resulted in the general reinstatement of instrumental performance (i.e., leverpressing was increased during both discriminative stimuli, regardless of which outcome they predicted during training). Given our findings, it seems likely that this lack of outcome selectivity in instrumental reinstatement was the result of the ambiguous O-R relationships present during training. These noncontingent outcome presentations were effective, however, in selectively reinstating conditioned magazine-approach responses to the stimulus that predicted that outcome during training, suggesting that Pavlovian reinstatement can be mediated by an S-O association. Therefore, it appears that a noncontingently delivered outcome can influence response selection through multiple routes.

AUTHOR NOTE

The present work was supported by National Institute of Mental Health Grant 56446. Correspondence concerning this article should be addressed to S. B. Ostlund, Department of Psychology, University of California, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563 (e-mail: sostlund@ucla.edu).

REFERENCES

- ADAMS, C. D., & DICKINSON, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **33B**, 109-121.
- ASRATYAN, E. A. (1974). Conditioned reflex theory and motivational behavior. *Acta Neurobiologiae Experimentalis*, **34**, 15-31.
- BALLEINE, B. W., & DICKINSON, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Psychopharmacology*, **37**, 407-419.
- BOLLES, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review*, **79**, 394-409.
- CAPALDI, E. J. (1967). Sequential hypothesis of instrumental learning. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 67-156). New York: Academic Press.
- COLWILL, R. M. (1994). Associative representations of instrumental contingencies. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 31, pp. 1-72). New York: Academic Press.
- COLWILL, R. M., & RESCORLA, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 120-132.
- COLWILL, R. M., & RESCORLA, R. A. (1986). Associative structures in instrumental conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 55-104). New York: Academic Press.
- 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.
- DELAMATER, A. R. (1997). Selective reinstatement of stimulus-outcome associations. *Animal Learning & Behavior*, **25**, 400-412.
- DELAMATER, A. R., LOLORDO, V. M., & SOSA, W. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Learning & Behavior*, **31**, 393-402.
- FRANKS, G. J., & LATTAL, K. A. (1976). Antecedent reinforcement schedule training and operant response reinstatement in rats. *Animal Learning & Behavior*, **4**, 374-378.
- HOLLAND, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, **30**, 104-117.
- KRUSE, J. M., OVERMIER, J. B., KONZ, W. A., & ROKKE, E. (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learning & Motivation*, **14**, 165-181.
- LERI, F., & STEWART, J. (2001). Drug-induced reinstatement to heroin and cocaine seeking: A rodent model of relapse in polydrug use. *Experimental & Clinical Psychopharmacology*, **9**, 297-306.
- MACKINTOSH, N. J., & DICKINSON, A. (1979). Instrumental (Type II) conditioning. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation* (pp. 143-167). Hillsdale, NJ: Erlbaum.
- REID, R. L. (1958). The role of the reinforcer as a stimulus. *British Journal of Psychology*, **49**, 202-209.
- RESCORLA, R. A. (1992). Response-outcome versus outcome-response associations in instrumental learning. *Animal Learning & Behavior*, **20**, 223-232.
- RESCORLA, R. A. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior*, **22**, 27-33.
- RESCORLA, R. A., & COLWILL, R. M. (1989). Associations with anticipated and obtained outcomes in instrumental learning. *Animal Learning & Behavior*, **17**, 291-303.
- RESCORLA, R. A., & SKUCY, J. C. (1969). The effect of response-independent reinforcers during extinction. *Journal of Comparative & Physiological Psychology*, **67**, 381-389.
- TRAPOLD, M. A., & OVERMIER, J. B. (1972). The second learning process in instrumental learning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 427-452). New York: Appleton-Century-Crofts.

(Manuscript received October 25, 2004;
revision accepted for publication November 14, 2006.)