

## **Effects of Outcome Devaluation on the Performance of a Heterogeneous Instrumental Chain**

**Bernard W. Balleine,**  
*University of California, Los Angeles, U.S.A.*

**Concepción Paredes-Olay, and**  
*University of Jaen, Spain*

**Anthony Dickinson**  
*University of Cambridge, U.K.*

Two experiments investigated the effects of outcome devaluation induced by conditioned taste aversion on the performance of the components of a heterogeneous chain of instrumental actions. In Experiment 1, thirsty rats were trained to perform two actions, R1 and R2, i.e., chain pulling and lever pressing counterbalanced, in sequence to gain access to a sucrose outcome in a single session; i.e.,  $R1 \rightarrow R2 \rightarrow O$ . Immediately after this session or after a delay the rats were injected with lithium chloride and given an extinction test on the two actions the next day. Although the immediate and delayed groups did not differ in the incidence of R1 on test, the immediate group reduced their performance of R2 relative to the delayed group. Experiment 2 assessed the effect of incentive learning after outcome devaluation. All rats were given an injection of lithium chloride immediately after training on the heterogeneous chain for sucrose reward after which half of the rats were reexposed to the sucrose whereas the remainder were reexposed to water prior to the extinction test. Although reexposure had no effect on the test performance of the R2 component in the chain, it significantly reduced performance of R1. These data are consistent with previous claims that responses proximal to outcome delivery are influenced by a Pavlovian incentive process whereas those more distal are controlled by the experienced incentive value of the outcome; i.e., by an instrumental incentive process.

In recent years there has been renewed interest in incentive processes and the role that they play in encoding emotional events (Flaherty, 1996; Dickinson & Balleine, 2002; Balleine, 2001; 2004, for reviews). For example, at Rutgers University, Flaherty and his colleagues have, over many years, carefully documented the effect of contrasting emotional expectations on incentive value by changing the motivational properties of sugar and other commodities across the phases of their experiments (e.g., Flaherty, 1996; Flaherty, Turovsky, & Krauss, 1994; Flaherty, Greenwood, Martin & Leszczuk, 1998). We have also been interested in this issue, particularly in the processes through which the incentive value of the reinforcer or outcome is modified in instrumental conditioning. Like Flaherty, we have found that changes in the incentive value of events like foods and fluids are largely a product of encoding relative changes in value on the basis of direct consummatory experience, a process we have called incentive learning (Dickinson & Balleine, 1994).

This research was conducted at the University of Cambridge with the support of grants from the European Commission BIOMED 2 (PL 950608) the Medical Research Council program (G9537855) and from the Ministerio de Educación y Cultura de España (BFPI-AP93). The report of these experiments was supported by the National Institute of Mental Health, grant #56446. Correspondence concerning this article should be addressed to Bernard W. Balleine, Department of Psychology, UCLA, Box 951563, Los Angeles, CA, 90095-1563, U.S.A. (balleine@psych.ucla.edu).

The effect of outcome devaluation produced by conditioned taste aversion provides a particularly powerful example of the influence of changes in incentive value on instrumental performance. In the first report of this effect, Adams and Dickinson (1981) trained hungry rats to lever press for sucrose pellets after which the incentive value of the pellets was altered by pairing their consumption with illness induced by an injection of lithium chloride (LiCl). Over the course of several pairings, the consumption of the pellets was reduced relative to a group in which pellet consumption and illness were unpaired. When the rats' instrumental performance was tested in extinction, lever pressing in the paired group was also reduced relative to the unpaired group. More recently, we have evaluated two contrasting accounts of the way that illness affects the value of the sucrose in this type of experiment (Balleine & Dickinson, 1991). On one account, changes in value reflect a signaling process; i.e., the effect of pairing the sucrose with illness is that the sucrose then acts as a signal for forthcoming illness. The alternative hypothesis, derived from Garcia's (1989) account of taste aversion learning, is that devaluation requires two steps: first the formation of a latent association between the representation of the sucrose and the processes activated by illness and, second, the experience of the aversive feedback now produced on consummatory contact with the sucrose due to the latent association. It is this latter experience that provides the opportunity for incentive learning; i.e., to encode the changed incentive value of the sucrose.

We recognized that, on the signaling account, a single pairing of the sucrose with illness, if it is effective, should also be sufficient to produce an outcome devaluation effect in subsequent instrumental performance whereas on the incentive learning account this pairing should only be effective after the rats had been reexposed to the sucrose after the pairing. In a test of these predictions we trained thirsty rats to lever press for sucrose in a single session after which they were given an injection of LiCl either immediately or after a delay. If this single pairing of sucrose with illness was effective, we should anticipate that the rats would show a devaluation effect and reduce their lever press performance in the subsequent extinction test whether or not they were reexposed to the sucrose after the illness. On the incentive learning account, however, this devaluation effect should only be observed in rats allowed to recontact the sucrose and so learn through direct experience that it is now aversive prior to the test. In a number of experiments of this kind we have found consistent evidence to support the incentive learning account. Only rats that were reexposed to the sucrose after it was paired with illness showed a devaluation effect in a subsequent extinction test (Balleine & Dickinson, 1991; 1992; Lopez, Balleine & Dickinson, 1992).

We concluded from these studies that incentive learning is necessary to show outcome devaluation effects in instrumental performance. Subsequently, however, Rescorla (1992, 1994) provided evidence that questioned the necessity of incentive learning. Rescorla trained hungry rats to press a lever and pull a chain with one action earning food pellets and the other a sucrose solution. The rats were then given a single pairing between one of the outcomes and illness before being given a choice extinction test on the levers and chains. Although no incentive learning was given, Rescorla found a significant reduction in performance on the action that, in training, had delivered the outcome paired with illness (see Parades-Olay & Lopez, 2000, 2002).

Although Rescorla's (1992, 1994) finding contrasts with the effects that we reported for lever press performance, it is, in fact, similar to what we have found on other measures, notably the response of operating a panel to gain access to the magazine. In our studies (Balleine & Dickinson, 1991, 1992; Lopez, Balleine & Dickinson, 1992), we consistently found evidence of a reduction in magazine entry via the panel after the outcome was devalued by taste aversion conditioning whether or not the rats were given incentive learning. Similarly, Balleine (1992) found that shifts in primary motivational state, while having no direct impact on lever pressing in the absence of incentive learning, directly modulated the magazine-panel response. Indeed, Balleine, Garner, Gonzalez, and Dickinson (1995) reported that removing the magazine panel rendered lever pressing directly sensitive to shift in primary motivational state. It is notable that access to the magazine did not require a panel response in Rescorla's (1992, 1994) apparatus.

It appears, therefore, that the contribution of incentive learning to the effect of outcome revaluation on instrumental responding depends on the location of the response within the heterogeneous behavioral chain required for reinforcement. If the response is a distal link in the chain then incentive learning is required for outcome revaluation to affect responding whereas responses proximal to reward are more likely to be directly modulated by the revaluation treatment. Balleine et al. (1995) examined this claim directly by training rats on an explicit chained schedule to perform one then another instrumental response (e.g., lever pressing followed by chain pulling or vice versa) for a food pellet reward in the absence of a magazine panel. When the food pellets were devalued by satiating the rats on their maintenance diet the rats reduced their performance of the proximal but not the distal response. Furthermore, when the rats were reexposed to the pellets after being satiated and thereby given the opportunity for incentive learning, performance of the distal response was reduced relative to rats not given incentive learning (Balleine et al., 1995). The proximal response of the chain was not further affected by incentive learning (see Corbit & Balleine, 2003).

These studies suggest that, after shifts in primary motivation, responses proximal to reward are more immediately influenced by outcome revaluation treatments than those more distal to reward with the latter only modified by incentive learning. The aim of the current series of experiments was to examine whether the same is true when the value of the instrumental outcome is modified by conditioned taste aversion procedures.

### **Experiment 1**

In Experiment 1, rats were trained to perform two actions, lever pressing and chain pulling, such that the delivery of the outcome contingent upon the performance of the second or terminal action (R2) depended upon prior execution of the first or initial action (R1). This schedule was designed to model the contingency operating in our previous experiments (Balleine & Dickinson, 1991, 1992), where lever pressing acted as R1 and pushing the magazine panel acted as R2, in similar fashion to our previous report (Balleine et al., 1995).

Thirsty rats were initially trained to perform R1 then R2 for a water reward. For half of the rats, chain pulling served as R1 and lever pressing as R2, whereas the remainder received the opposite ordering of actions. After this training, per-

formance on the chained schedule was partially extinguished before all of the rats were given a single training session in which performance on this schedule was rewarded with a 20% sucrose solution. Immediately after this session, half of the rats were given an injection of LiCl, Group IMM, whereas the remainder were given the injection after a 6-h delay, Group DEL. Balleine and Dickinson (1991, 1992) found that this procedure devalued the reinforcing properties of the sucrose in Group IMM but not in Group DEL. If, as was argued with panel pushing to enter the magazine, the action proximal to outcome delivery (i.e., R2) is more sensitive to outcome devaluation, the performance of R2 in Group IMM should be reduced relative to Group DEL when subsequently tested in extinction. Further, if more distal actions are relatively insensitive to outcome devaluation, as was argued to be the case with lever pressing prior to magazine entry, the test performance of R1 should not differ between groups.

### **Method**

**Subjects and Apparatus.** Sixteen naive adult male hooded Lister rats were housed in squads of four. They were maintained on Rat and Mouse No. 1 (Modified) low-protein, high-fiber, expanded pellets (Special Diet Services, UK). The animals were trained and tested in four Campden Instruments (Manchester, UK) operant chambers housed in sound- and light-resistant shells. Each chamber was equipped with a dipper that could deliver either 0.05 ml of either tap water or a 20% sucrose solution, as appropriate, into a recessed magazine located in the center of the front wall. The magazines were fitted with transparent perspex magazine panels that were fixed in a permanently open position so that access to the reinforcer did not require a magazine-panel response. Each chamber was also equipped with a retractable lever located to the right of the magazine, and a chain that could be lowered through the roof from a microswitch so that it hung to the left of the magazine 3.5 cm from the front wall. The lever and chain were positioned symmetrically to the right- and left-hand sides of the magazine, respectively. Each chamber was illuminated by a 3-W, 24-V house light mounted in the center of the front wall above the magazine. A BBC microcomputer, equipped with the SPIDER extension for on-line control (Paul Fray Ltd., Cambridge, UK), controlled the equipment and recorded the lever presses and chain pulls. Training and testing took place in the light cycle between 9:00 and 16:00 h each day.

**Water Pretraining.** All rats were maintained on a 22.5-h water deprivation schedule throughout training with free access to their maintenance diet allowed in the home cage and 1.5 h of access to water given each day in the home cage at least 2 h after the daily training session.

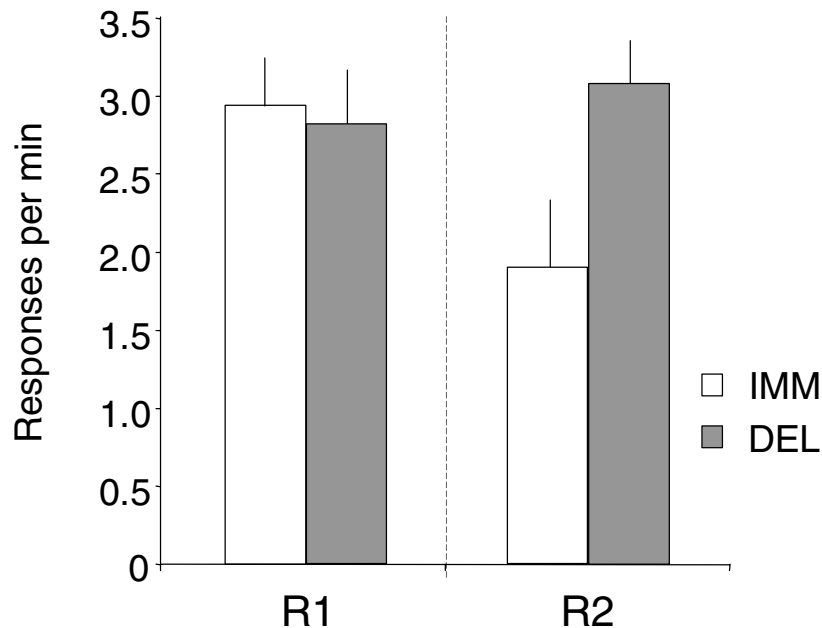
The rats initially received two sessions of magazine training in each of which 30 water reinforcers were delivered on a variable time (VT) 60-s schedule with the levers and chains retracted. After magazine training, all rats received 15 sessions of training on the heterogeneous chain of instrumental actions for a water reward. The first response in this chain was designated R1 and the second R2. For half of the rats R1 was lever pressing and R2 chain pulling. For the remaining rats, R1 was chain pulling and R2 lever pressing. In the first session all animals were trained to perform R2 with the water outcome delivered on a continuous reinforcement schedule until 30 deliveries had been earned with the R1 manipulandum retracted. After this session, the R1 manipulandum was introduced and the water was available on a variable interval (VI) schedule. Scheduled water deliveries were contingent on the performance of R2 given that R1 had been executed at least once after the interval schedule determined that the water outcome was available. No other constraint was placed on the relation between R1 and R2. The parameter for the VI schedule was 2 s for the first 5 sessions of training and was then incremented to 15 s for the next five and finally to 30 s for the final five sessions of training. Each training session began with the onset of the house lights and continued until 30 water deliveries had been earned at which point the house lights were turned off and the session terminated. The day after the final training session all rats were given a 30-min extinction session with the levers and chains in preparation for the sucrose training session.

**Sucrose Training.** The day after the extinction session, all of the rats were given sucrose training conducted in a single 30-min session in which performance of the chain was rewarded with

0.05 ml of 20% sucrose solution delivered on an VI-10-s schedule. During this session we also assessed sensitivity to the chain contingency. To this end, two conditional probabilities were examined: the probability of performing R2 in each second after R1 had been performed, i.e.,  $p(R1 \rightarrow R2)$ , and the probability of performing R1 in each second after R2 had been performed, i.e.,  $p(R2 \rightarrow R1)$ . To calculate  $p(R1 \rightarrow R2)$ , for each rat the number of R2s performed in each successive 1-s bin after the performance of R1 was totaled for each bin throughout the training session and divided by the number of times performance in each bin was sampled during the session (i.e., the number of opportunities for executing R2 in each of the 1-s bins after the performance of R1). A comparable assessment of the probability of R1 following R2 was also conducted.

Immediately after this session half of the rats in each R1-R2 counterbalancing condition, matched for their performance during the pretraining phase, were given an injection of LiCl (0.15 M; 20 ml/kg, i.p.) and constituted Group IMM ( $n = 8$ ). The remaining rats received the injection of LiCl after a delay of 6 h and constituted Group DEL ( $n = 8$ ). The daily water ration was given to all of the rats at least 2-h after the injection was given to Group DEL.

**Extinction and Reinforcement Tests.** The day after sucrose training, the rats were given an extinction test on the levers and chains. The session began with the onset of the house light and terminated with its offset after 30 min. No outcomes were delivered during this session. The next day, reacquisition was conducted on the chain with water delivered on a VI-30-s schedule. A reinforced test was conducted the following day in a session in which the sucrose was earned by performance on the chain on a VI-30-s schedule. This session was otherwise the same as the extinction test and lasted 30 min.

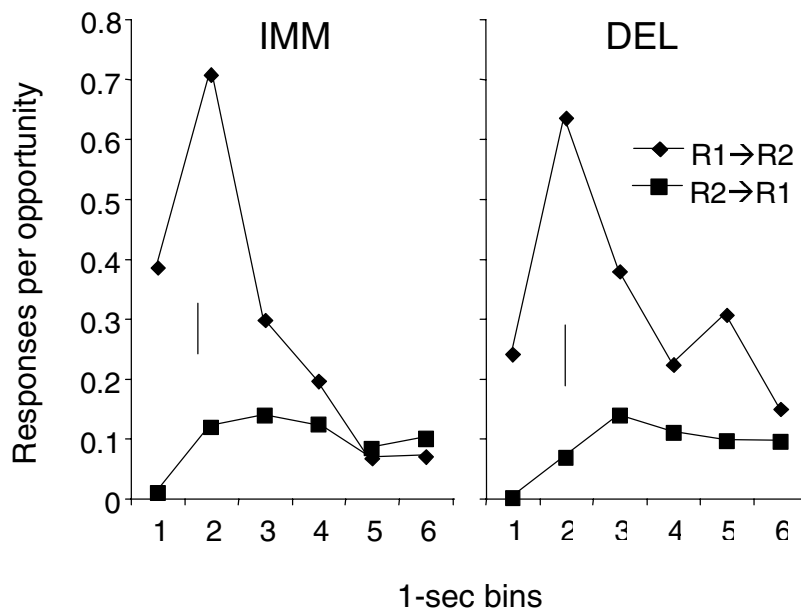


**Figure 1.** Experiment 1. Mean responses per minute performed in the extinction test of Experiment 1. Performance of the actions distal (R1; left panel) and proximal (R2; right panel) to outcome delivery are shown separately for rats in Group IMM and Group DEL. Bars =  $\pm 1$  SEM.

### Results and Discussion

The results from the extinction test are presented in Figure 1. Generally, it appeared that, in Group IMM, the rats performed fewer of R2 than of R1. In Group DEL, by contrast, animals performed, if anything, more R2 than R1. Further, although the performance of R1 was comparable in Group IMM and Group DEL, rats in Group IMM performed fewer of R2 than Group DEL. This pattern of results

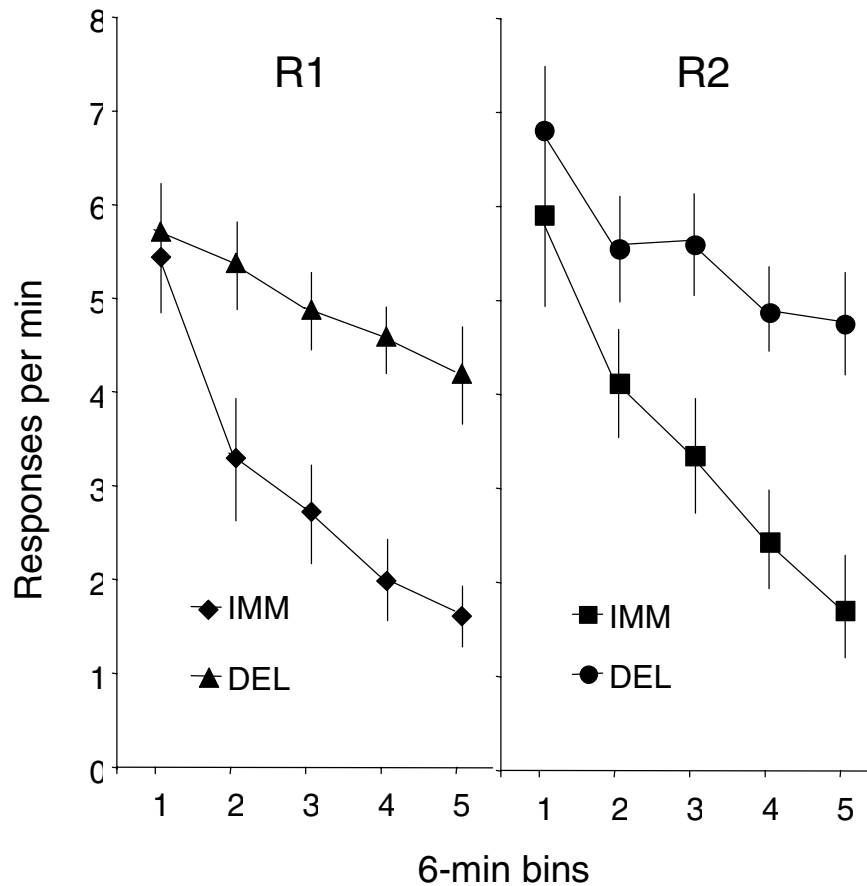
was confirmed by the statistical analysis when reliability was assessed against a Type I error of 0.05. A two-way mixed analysis of variance was conducted on the total lever presses and chain pulls in the extinction test with a between-subjects factor of group, separating performance in Group IMM and Group DEL, and a within-subjects factor of action, separating performance of the action designated R1 from the performance of R2 in the chain. This analysis found no effect of action ( $F < 1$ ) but revealed a significant effect of group,  $F(1, 14) = 6.1$ , and a significant Group  $\times$  Action interaction,  $F(1, 14) = 5.5$ . Simple main effects analyses conducted on the interaction revealed that, although the effect of action was not reliable in Group DEL ( $F = 1.6$ ), it was significant in Group IMM,  $F(1, 14) = 7.6$ . Further, whereas there was an effect of group on R2,  $F(1, 14) = 11.9$ , there was no effect of this factor on R1 ( $F < 1$ ).



**Figure 2.** Experiment 1. The mean number of actions per opportunity for rats in Group IMM, left panel, and Group DEL, right panel, during the sucrose training session. The two curves in each panel represent the probability of performing R2 in each 1-s bin after the performance of R1 (i.e., R1 → R2 - diamonds) and the probability of performing R1 in each 1-s bin after the performance of R2 (i.e., R2 → R1 - squares). Bars =  $\pm 1$  standard error of the difference of the mean (SED) on the within-subjects variable.

These effects emerged on test and were not present in the sucrose training session. A comparable analysis conducted on the rate of lever pressing and chain pulling in that session revealed no effects of group or of action nor an interaction between these factors ( $F_s < 1$ ). The mean performance of R1 and R2 per minute, respectively, in the sucrose training session were; in Group IMM: 5.8 ( $\pm 0.8$ ) and 6.9 ( $\pm 1.1$ ); and in Group DEL: 5.5 ( $\pm 0.9$ ) and 6.5 ( $\pm 0.7$ ). Additional analysis of performance on the components of the chain during that session suggested that the rats in both Groups IMM and DEL were sensitive to the chain contingency. Figure 2 displays the probability of performing one action as a function of time since performance of the other action for both groups. It is clear from this figure that the likelihood of performing R2 was at a maximum in the first two seconds after an

execution of R1 in both groups. This relationship was not reciprocal, however—performance of R1 was not systematically affected by whether or not the rats had emitted R2 in the preceding 6 s. An analysis was conducted on these data using a between-subjects factor of group and within-subjects factors of action order, separating the probability of performing R2 after R1 from that of R1 after R2, and of time bin. This analysis found no effect of group ( $F < 1$ ) but found effects of action order,  $F(1, 14) = 31.3$ , of time bin,  $F(5, 70) = 4.0$ , and an Action order  $\times$  Time Bin interaction,  $F(5, 70) = 15.9$ . No other interactions were significant (largest  $F(9, 126) = 1.8$ ). Simple main effects analyses conducted on the significant interaction revealed that there was an effect of time bin on the performance of R2 after R1,  $F(5, 70) = 12.1$ , but not on the performance of R1 after R2 ( $F < 1$ ). In addition, a significant effect of action order was found in bin 1,  $F(1, 14) = 18.7$ , and in bin 2,  $F(1, 14) = 46.4$ , but not in any other bin (largest  $F(1, 14) = 2.9$ ).



**Figure 3.** Experiment 1. Mean responses per minute during the reacquisition session conducted with the sucrose reward. Responding on R1 (left panel) and R2 (right panel) of the heterogeneous chain is presented separately in 6-min bins across the 30 min session for Group IMM (R1 = diamonds; R2 = squares) and Group DEL (R1 = triangles; R2 = circles). Bars =  $\pm 1$  SEM.

The results of this experiment suggest that the immediate effects of outcome devaluation differ on the components of a heterogeneous chain; the action proximal to the delivery of the outcome was found to be more sensitive to the effects of

devaluation than the more distal action. It is not clear, however, whether this difference reflects a differential impact of devaluation or whether the control of performance by the representation of the outcome differs in the case of the two actions. If it is the former, then delivering the devalued outcome as a consequence of performance on the chain should be less effective in producing a change in the performance of R1 than in the performance of R2. In contrast to this prediction, however, the results of the reinforcement test session conducted with the sucrose outcome, presented in Figure 3, suggest that both R1 and R2 were similarly sensitive to the effects of devaluation when the devalued outcome was actually delivered. Although responding was similar in both Group IMM and DEL at the outset, Group IMM quickly reduced the performance of both R1 and R2 relative to the performance of these responses in Group DEL as the test session continued. An analysis conducted on these data revealed a main effect of group,  $F(1, 14) = 13.2$ , but neither an effect of action nor a Group  $\times$  Action interaction (largest  $F(1, 14) = 2.2$ ). Furthermore, there was a main effect of bin,  $F(4, 56) = 28.4$ , and a significant Group  $\times$  Bin interaction,  $F(4, 56) = 15.7$ , confirming that the difference in the performance on the two responses in Group IMM and Group DEL emerged over the course of test session.

The results of this experiment confirm the suggestion that, in a chain of actions composed of lever pressing and chain pulling, the action proximal to the delivery of the outcome is more sensitive to outcome devaluation produced by taste aversion conditioning than the more distal action. Indeed, it appears from the results of the extinction test that the first action in the chain of responses (R1) was relatively unaffected by outcome devaluation; the rats performed an equivalent number of R1 in the extinction test whether they were given the injection of LiCl immediately or after a delay. Importantly, this pattern of results parallels that previously observed for lever pressing and magazine entry (e.g., Balleine & Dickinson, 1991, 1992) suggesting that the previously reported difference in the sensitivity of these actions to outcome devaluation could be due to their relative locations within a behavioral chain.

What remains unclear, however, is whether the first link of the present heterogeneous chain is sensitive to an incentive learning treatment. If R1 plays the same functional role as lever pressing did in our previous studies (Balleine & Dickinson, 1991, 1992), then exposure to the devalued outcome prior to the extinction test should be predicted to reduce the performance of R1 on test. This prediction was assessed in Experiment 2.

## Experiment 2

In Experiment 2, thirsty rats were trained on the chain first for water and then for sucrose as described in Experiment 1. Immediately after the sucrose training session, all of the rats were given an injection of LiCl. The next day, half of the rats were reexposed to the sucrose, Group SUC, whereas the remainder were reexposed to water, Group H2O, prior to an extinction test conducted on the levers and the chains.

The critical prediction concerns the effect of reexposure to the sucrose on performance of the first action of the chain, R1. In Group H2O we expected to replicate the pattern of performance on R1 and R2 observed in Group IMM in Ex-

periment 1; i.e., that the animals should perform R1 more frequently than R2. If R1 is sensitive to incentive learning, however, performance of that action should be reduced in Group SUC relative to Group H2O.

### **Method**

**Subjects and Apparatus.** Sixteen naive adult male hooded Lister rats were housed and maintained as described in Experiment 1 and trained in the same operant chambers.

**Water Pretraining.** The maintenance conditions and procedures for initial training on the chain were identical to those described in Experiment 1 except that the rats received only 12 sessions of training on the heterogeneous chain of instrumental actions for a water reward; i.e., R1 → R2 → H2O. For half of the rats in Experiment 2, R1 was lever pressing and R2 was chain pulling. For the remaining rats, R1 was chain pulling and R2 was lever pressing. The parameter for the VI schedule was 2 s for the first four sessions of training and was then incremented to 15 s for the next four and finally to 30 s for the final four sessions of training. As in Experiment 1, the day after the final training session all rats were given a 30-min extinction session on the levers and chains in preparation for the sucrose training session.

**Sucrose Training.** All of the rats were given sucrose training as described in Experiment 1. This session was 30 min in duration and performance of the chain was rewarded with 0.05 ml of 20% sucrose solution delivered on a VI-10 s schedule. Again sensitivity to the chain contingency was also assessed during this session as described in Experiment 1. Immediately after this training session all of the rats were given an injection of LiCl (0.15 M; 20 ml/kg, i.p.). The daily water ration was given to all of the rats at least 2-h after the injection.

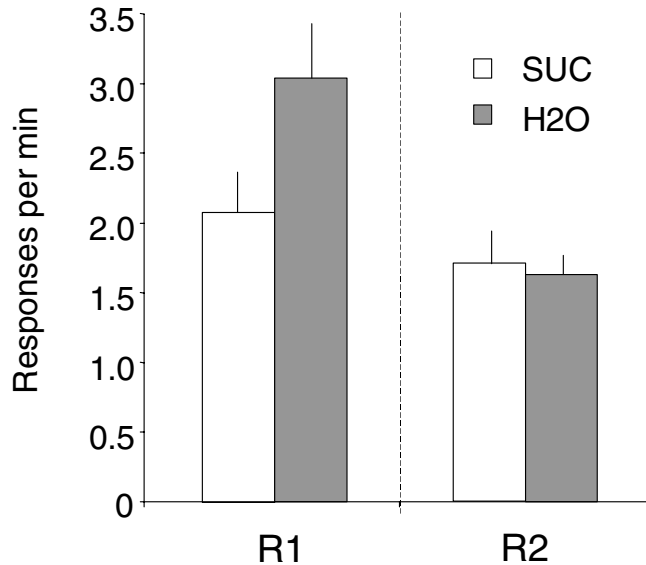
**Reexposure.** The next day the rats were given a single reexposure session in the operant chambers with the levers and chains removed. For half of the rats, 0.05 ml of the sucrose solution was delivered into the recessed magazine on a VT-60 s schedule until thirty deliveries had been given. These rats constituted Group SUC ( $n = 8$ ). The remaining rats were placed in the operant chambers but were not reexposed to sucrose. Instead these rats were reexposed to 0.05 ml of tap water again delivered on a VT-60 s schedule until 30 deliveries had been given. This group constituted Group H2O ( $n = 8$ ).

**Extinction and Reinforcement Tests.** The day after reexposure, the rats were given a 30-min extinction test on the levers and chains. No outcomes were delivered during this session. The next day, reacquisition was conducted on the heterogeneous chain with water delivered on a VI 30-s schedule. A reinforced test was conducted the next day in which the sucrose was earned by performance on the chain on a VI 30-s schedule. This session again lasted 30 min.

### **Results and Discussion**

The results of the extinction test are presented in Figure 4. The effect of conditioning a taste aversion to the sucrose in Group H2O appeared to be very similar to that observed in Group IMM in Experiment 1; i.e., the rats performed fewer of R2 than of R1. Reexposure to sucrose prior to the test produced a reduction in R1 such that R1 and R2 were performed at a more comparable rate in Group SUC. Furthermore, the performance of R2 was similar in the two groups, whereas Group H2O performed more of R1 than Group SUC. This pattern of results was confirmed by the statistical analysis. A two-way mixed analysis of variance was conducted on the total lever presses and chain pulls in the extinction test with a between-subjects factor of group, separating performance in Group SUC and Group H2O, and a within-subjects factor of action, separating performance of the action designated R1 from the performance of R2 in the chain. This analysis found no effect of group, ( $F(1, 14) = 1.3$ ), but revealed a significant effect of ac-

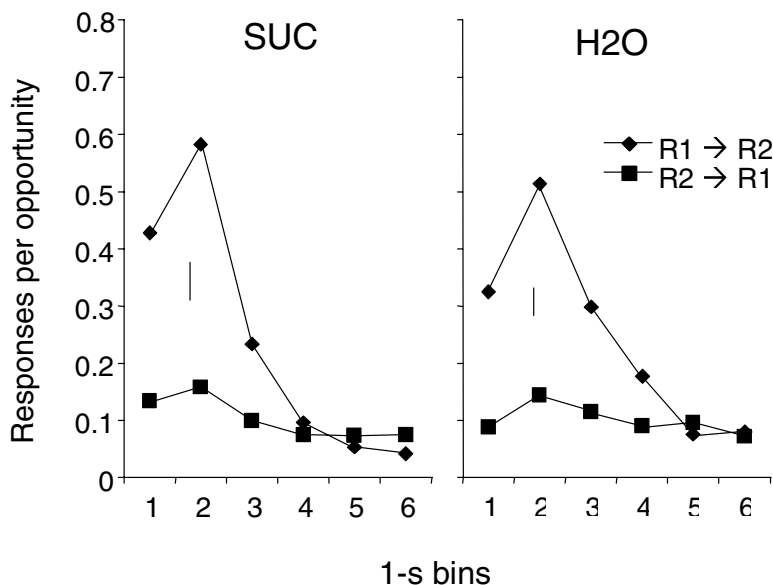
tion,  $F(1, 14) = 23.8$ , and a significant Group  $\times$  Action interaction,  $F(1, 14) = 10.1$ . Simple main effects analyses conducted on the interaction revealed that, although the effect of action was not reliable in Group SUC,  $F = 1.5$ , it was significant in Group H2O,  $F(1, 14) = 32.5$ . Furthermore, whereas there was an effect of group on the performance of R1,  $F(1, 14) = 7.0$ , there was no effect of this factor on the performance of R2,  $F < 1$ .



**Figure 4.** Experiment 2. Mean responses per minute performed in the extinction test of Experiment 2. Performance of the actions distal (R1; left panel) and proximal (R2; right panel) to outcome delivery are shown separately for rats in Group SUC and Group H2O. Bars =  $\pm 1$  SEM.

Again, these effects emerged on test and were not present in the sucrose training session. A comparable analysis conducted on the rate of lever pressing and chain pulling in that session revealed no effects of group or of action nor an interaction between these factors,  $F_s < 1$ . The mean performance of R1 and R2 per minute, respectively, in the sucrose training session were; in Group SUC: 4.6 ( $\pm 0.5$ ) and 5.7 ( $\pm 0.7$ ); and in Group DEL: 4.3 ( $\pm 0.4$ ) and 5.2 ( $\pm 0.6$ ). As in Experiment 1, additional analysis of performance on the components of the chain during the sucrose training session suggested that the rats in both Groups SUC and H2O were sensitive to the chain contingency and to a similar degree. As shown in Figure 5, the likelihood of performing R2 was at a maximum in the first 1 or 2 s after an execution of R1 in both groups. In contrast, performance of R1 was not systematically affected by whether or not the rats had emitted R2 in the preceding 6 s. An analysis was conducted on these data using a between-subjects factor of group and within-subjects factors of action order, separating the probability of performing R2 after R1 from that of R1 after R2, and of time bin. This analysis found no effect of group,  $F < 1$ , but found effects of action order,  $F(1, 14) = 10.5$ , of time bin,  $F(5, 70) = 17.6$ , and an Action Order  $\times$  Time Bin interaction,  $F(5, 70) = 7.5$ . No other interactions were significant, all  $F_s < 1$ . Simple main effects analyses conducted on the significant interaction revealed that there was an effect of time bin on the performance of R2 after R1,  $F(5, 70) = 13.6$ , but not on the performance of R1 after R2,  $F < 1$ . In addition, a significant effect of action order was found in bin 1,

$F(1, 14) = 11.4$ , and in bin 2,  $F(1, 14) = 18.9$ , but not in any other bin, largest  $F(1, 14) = 3.7$ .

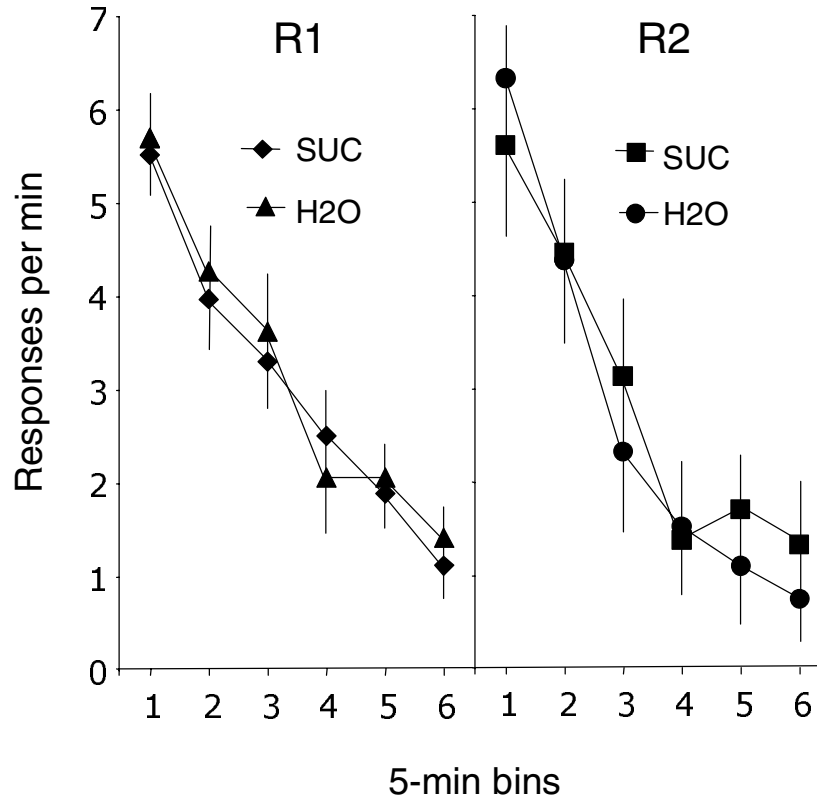


**Figure 5.** Experiment 2. The mean number of actions per opportunity for rats in Group SUC, left panel, and Group H2O, right panel, during the sucrose training session. The two curves in each panel represent the probability of performing R2 in each 1-s bin after the performance of R1 (i.e., R1 → R2 - diamonds) and the probability of performing R1 in each 1-s bin after the performance of R2 (i.e., R2 → R1 - squares). Bars =  $\pm 1$  SED on the within-subjects variable.

Although unlikely, it is at least logically possible that, using the current parameters, the group differences in the performance on R1 and R2 observed during the extinction test were the product of changes in the strength of the aversion conditioned to the sucrose. Perhaps reexposure to the H2O interfered with the conditioning of the sucrose aversion in Group H2O or, by acting as a reminder, perhaps reexposure to the sucrose produced a relative increase in the aversion in Group SUC. The results of the reinforcement test session conducted with the sucrose outcome, and presented in Figure 6, do not support either of these suggestions however. Rather they indicate that both R1 and R2 were similarly sensitive to the effects of devaluation when the devalued outcome was actually delivered and to a comparable degree in both groups. There was a significant main effect of bin,  $F(5,70) = 38.6$ , but no other main effects nor interactions were significant, largest  $F(5,70) = 2.1$  confirming that the sucrose aversion did not differ between groups.

As was found for Group IMM in Experiment 1, the rats given reexposure to water reduced their performance of the action proximal to outcome delivery relative to one more distal after devaluation of the sucrose. The important finding of the present study, however, was that the incentive learning treatment produced a reduction in the performance of the more distal action, R1, which was then performed at a comparable rate to the proximal action, R2. These data, along with those of Experiment 1, demonstrate that, when rats are trained on an explicit, heterogeneous instrumental chain, components of that chain are differentially sensitive to devaluation induced by conditioning a taste aversion to the instrumental

outcome, a difference that can be ameliorated by incentive learning.



**Figure 6.** Experiment 2. Mean responses per minute during the reacquisition session conducted with the sucrose reward. Responding on R1 (left panel) and R2 (right panel) of the heterogeneous chain is presented separately in 5-min bins across the 30 min session for Group SUC (R1 = diamonds; R2 = squares) and Group H2O (R1 = triangles; R2 = circles). Bars =  $\pm 1$  SEM.

### General Discussion

The results of the current experiments provide clear support for the argument that the sensitivity of instrumental actions to the effects of devaluation produced by conditioning a taste aversion to the instrumental outcome depends on the position of the action in a heterogeneous chain relative to reward delivery. Actions proximal to reward delivery were more immediately sensitive to devaluation than those more distal. Nevertheless, distal actions were influenced by the change in value but only after the rats had been provided with the opportunity for incentive learning.

These results provide a means of reconciling previously disparate reports from experiments assessing the effects of taste aversion-induced outcome devaluation on instrumental performance. They are also consistent with a number of previous reports and further our understanding of the incentive processes that contribute to instrumental performance. The finding in both experiments that taste aversion-induced outcome devaluation produced an immediate reduction in the subsequent performance of that action most proximal to reward delivery is similar to the previously reported effects of post-training shifts in primary motivational state

both on performance of a similar chain of actions but also on magazine entry via a panel (Balleine, 1992; Balleine et al., 1995). This immediate effect was also similar to that observed in previous studies after taste aversion-induced outcome devaluation on magazine entry (Balleine & Dickinson, 1991, 1992; Lopez et al. 1992) and has much in common with the immediate effect of outcome devaluation on instrumental choice performance observed in Rescorla's (1992; 1994) studies.

In contrast to the immediate effect of devaluation on the proximal action, the link in the chain of actions that was more distal to the delivery of the reward was not immediately sensitive to outcome devaluation. Again, this effect has been previously reported after a shift in primary motivation (Balleine et al. 1995) and in many ways is similar to the reported insensitivity of lever pressing to the effects of taste aversion-induced outcome devaluation. One simple explanation of this effect can be drawn from those that have been previously put forward to explain why second-order conditioned responses in Pavlovian conditioning are insensitive to US devaluation (e.g., Holland & Rescorla, 1975) i.e., that, unlike the proximal action, performance on the distal action is not determined by its relation to the outcome. Evidence against this suggestion is, however, provided by the results of Experiment 2. In that experiment it was found, as has been reported previously after taste aversion-induced outcome devaluation (Balleine & Dickinson, 1991, 1992; Lopez et al. 1992), that, after the pairing between sucrose and the injection of LiCl, reexposure to the sucrose outcome was sufficient to reduce performance of the distal action. Hence, although performance of the distal action was not immediately sensitive to outcome devaluation, it was affected by devaluation when the change in the incentive value of the outcome was made explicit through direct consummatory experience. At the very least, therefore, the association between the distal action and the outcome (i.e.,  $R1 \rightarrow O$ ) and that between the proximal action and the outcome (i.e.,  $R2 \rightarrow O$ ) cannot be identical.

It remains to be considered how these associations differ; whether the relationships of the distal and proximal actions with the outcome differ along a qualitative or, perhaps, a quantitative dimension or, alternatively, whether the functional effects of these associations reflect the operation of quite distinct incentive processes in instrumental conditioning. From the former perspective, for example, it is quite possible that, as an animal passes through an instrumental chain, its representation of the consequences of performance on that chain becomes more accurate; perhaps distal actions, requiring as they do the retrieval of a consequent still some temporal and, possibly, spatial distance away, are associated with a more diffuse outcome representation, whereas responses that are proximal to the outcome have associations with its more specific features. Hence, manipulations that change the value of the outcome might be expected to have their greatest impact on the responses more proximal to it; at least to the extent that these changes in value can be assumed to affect these specific features. Nevertheless, although this account can readily explain the differential effects of outcome devaluation on the proximal and distal elements of the chain in Experiment 1, it has some difficulty generating an account of the results of Experiment 2. It is simply unclear why, after the devaluation treatment, reexposure should render the representation of the outcome associated with the distal action more accurate, at least sufficiently so to produce an outcome devaluation effect on that action.

In contrast, as was argued by Balleine (2001), it is possible that, rather than

generating a more accurate representation of the outcome, reexposure acts to devalue those aspects of the outcome representation with which the distal action is associated. On this account, due to overshadowing produced by its greater contiguity with outcome delivery, the proximal action is more strongly associated with the most salient features of the outcome whereas the distal action is associated with more diffuse or less salient features. If, as seems likely, the injection of LiCl also produces a greater immediate change in the value of the more salient than the less salient features of the outcome, it naturally follows that the proximal action should be more strongly affected than the distal action. Furthermore, if, during reexposure, the presentation of these now highly aversive, highly salient features contiguous with the less salient features allows the latter now to become aversive, then a ready explanation of why reexposure works to reduce performance on the distal action can be proposed that is entirely consistent with the direct effect of devaluation on the proximal action.

Rather than attempting to explain responding on the proximal and distal actions using a unified account, it may alternatively be proposed that the distal and proximal actions are both related to distinct outcome representations that are, themselves, under the control of distinct incentive processes. It is possible, for example, that the rate of performance of actions that are proximal to reward delivery at least partly reflects the contribution of the excitatory effects of Pavlovian stimulus-outcome associations; the sight of the instrumental manipulandum could, through association with the outcome, provide a source of excitation that bolsters performance on that action (Rescorla & Solomon, 1967). If pairing the outcome with illness removes this source of excitation, then it should also produce an immediate devaluation effect in any Pavlovian CR's and in the proximal action without the need for reexposure whilst leaving the performance of more distal actions unaffected. In contrast, due to overshadowing by stimuli related to the proximal response, performance of the distal action is not modulated by stimulus-outcome associations. Instead, as we have argued previously, this action is controlled by an instrumental incentive process that encodes the incentive value of instrumental outcomes (Dickinson & Balleine, 1993, 2002; Balleine, 2004, for reviews).

Some support for this perspective has been reported. For example, Corbit and Balleine (2003) found that, whereas performance of the proximal action in a heterogeneous chain was increased by the presentation of a separately trained Pavlovian excitator, the distal action was not affected by these stimuli. Furthermore, as was found in Experiment 2, Corbit and Balleine (2003) demonstrated that incentive learning was only effective in modifying performance on the distal component of the chain. A similar double dissociation of the influence of Pavlovian cues and of incentive learning on instrumental performance has also been reported after lesions of distinct regions of the nucleus accumbens. Corbit, Muir, and Balleine (2001) found that lesions of the accumbens shell abolished the influence of Pavlovian cues on instrumental performance but did not affect the sensitivity of performance to outcome devaluation, in this case induced by sensory-specific satiety. In contrast, lesion of the accumbens core abolished sensitivity to outcome devaluation but left the excitatory effect of Pavlovian cues on instrumental performance intact. These studies support the notion that Pavlovian and instrumental incentive processes independently influence performance and that, whereas the former exerts a stronger influence on the proximal action, it is instrumental incentive learning that controls

performance of actions more distal to reward.

Whatever the source of the differential influence of outcome devaluation on components of a heterogeneous chain turns out to be, the current results have established that the effect of outcome devaluation on those components is similar whether devaluation is accomplished by a shift in primary motivation or by conditioning a taste aversion to the outcome. As such, these results join others in pointing to the general importance of incentive processes in the operation of the systems that control adaptive responding. Clearly the multistage processing of events proposed by Flaherty and colleagues (e.g., Flaherty et al., 1998) to underlie the determination of relative changes in incentive value has much in common with the nature of instrumental incentive learning, through which post-training changes in incentive value are also encoded. Flaherty et al. (1998) proposed that an early aspect of the reaction to reward reduction is searching (Flaherty, 1996) whereas the emotional components associated with changes in incentive value occur later during subsequent exposure to the rewarding event after its initial shift in value, something that is entirely consistent with the basis for the changes in incentive value that we have found support outcome devaluation effects in instrumental conditioning.

## References

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **33B**, 109-122.
- Balleine, B. W. (1992). The role of incentive learning in instrumental performance following shifts in primary motivation. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 236-250.
- Balleine, B. W. (2001). Incentive processes in instrumental conditioning. In R. Mowrer & S. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 307-366). Hillsdale, NJ: Erlbaum
- Balleine, B. W. (2004). Incentive Behavior. In: I.Q. Whishaw and B. Kolb (Eds). *The behavior of the laboratory rat: A handbook with tests* (pp. 436-446). Oxford, UK: Oxford University Press.
- Balleine, B. W., & Dickinson, A. (1991). Instrumental performance following reinforcer devaluation depends on incentive learning. *Quarterly Journal of Experimental Psychology*, **43B**, 279-296.
- Balleine, B. W., & Dickinson, A. (1992). Signalling and incentive processes in instrumental reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **45B**, 285-301.
- Balleine, B. W., Garner, C., Gonzalez, F. & Dickinson, A. (1995). Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*, **21**, 203-217.
- Corbit, L. H., & Balleine, B. W. (2003) Pavlovian and instrumental incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *Journal of Experimental Psychology: Animal Behavior Processes*, **29**, 99-106.
- Corbit, L., Muir, J. & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence for a functional dissociation between accumbens core and shell. *Journal of Neuroscience*, **21**, 3251-3260
- Dickinson, A., & Balleine, B. W. (1993). Actions and responses: The dual psychology of behaviour. In N. Eilan, R. McCarthy, M. W. Brewer (Eds.), *Spatial representation* (pp. 277-293). Oxford, UK: Basil Blackwell.
- Dickinson, A., & Balleine, B. W. (1994). Motivational control of goal-directed action. *Animal Learning and Behavior*, **22**, 1-18.
- Dickinson, A., & Balleine, B. W. (2002). The role of learning in motivation. In C. R. Gallistel (Ed.), *Learning, motivation and emotion, Volume 3 of Steven's Handbook of Experimental Psychology, Third Edition* (pp. 497-533). New York: John Wiley & Sons.
- Flaherty, C. F. (1996). *Incentive relativity*. New York, NY: Cambridge University Press.

Flaherty, C. F., Turovsky, J., & Krauss, K. L. (1994). Relative hedonic value modulates anticipatory contrast. *Physiology and Behavior*, **55**, 1047-1054.

Flaherty, C. F., Greenwood, A., Martin, J., & Leszczuk, M. (1998). Relationship of negative contrast to animal models of fear and anxiety. *Animal Learning and Behavior*, **26**, 397-407.

Garcia, J. (1989). Food for Tolman: Cognition and cathexis in concert. In T. Archer & L-G. Nilsson (Eds.), *Aversion, avoidance and anxiety* (pp. 45-85). Hillsdale, N.J.: Erlbaum.

Holland, P. C., & Rescorla, R. A. (1975). The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, **1**, 355-363.

Lopez, M., Balleine, B. W., & Dickinson, A. (1992). Incentive learning following reinforcer devaluation is not conditional upon the motivational state during re-exposure. *Quarterly Journal of Experimental Psychology*, 1992, **45B**, 265-284.

Paredes-Olay, C., & López, M. (2000). Comparison of the instrumental reinforcer devaluation effect in two strain of rats (Wistar and Lister). *Behavioural Processes*, **50**, 165-169.

Paredes-Olay, C., & López, M. (2002). Lithium-induced outcome devaluation in instrumental conditioning: Dose-effect analysis. *Physiology and Behavior*, **75**, 603-609.

Rescorla, R. A. (1992). Depression of an instrumental response by a single devaluation of its outcome. *Quarterly Journal of Experimental Psychology*, **44B**, 123-136.

Rescorla, R. A. (1994). A note on depression of instrumental responding after one trial of outcome devaluation. *Quarterly Journal of Experimental Psychology*, **47B**, 27-37.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationship between Pavlovian conditioning and instrumental learning. *Psychological Review*, **74**, 151-182.

*Received March 3, 2005.*

*Revision received April 28, 2005.*

*Accepted May 3, 2005.*