Motivational control after extended instrumental training

A. DICKINSON, B. BALLEINE, and A. WATT University of Cambridge, Cambridge, England

F GONZALEZ

Universidad Complutense de Madrid, Madrid, Spain

and

R. A. BOAKES

University of Sydney, Sydney, New South Wales, Australia

Hungry rats were trained to press a lever for food pellets prior to an assessment of the effect of a shift in their motivational state on instrumental performance in extinction. The first study replicated the finding that a reduction in the level of food deprivation has no detectable effect on extinction performance unless the animals receive prior experience with the food pellets in the nondeprived state (Balleine, 1992; Balleine & Dickinson, 1994). When tested in the nondeprived state, only animals that were reexposed to the food pellets in this state between training and testing showed a reduction in the level of pressing during the extinction test relative to animals tested in the deprived state. The magnitude of this reexposure effect depended, however, on the amount of instrumental training. Following more extended instrumental training, extinction performance was unaffected by reexposure to the food pellets in the nondeprived state whether or not the animals were food deprived at the time of testing. A second study demonstrated that the resistance to the reexposure treatment engendered by overtraining was due to the animals' increased experience of the food pellets in the deprived state during training rather than to the more extensive exposure to the instrumental contingency. In contrast to the results of the first two experiments, however, a reliable reexposure effect was detected after overtraining in a final study, in which the animals were given greater reexposure to the food pellets in the nondeprived state.

Drive theory, developed largely within the framework of the stimulus—response/reinforcement account of instrumental conditioning (e.g., Hull, 1943), assumes that primary motivational states, such as hunger and thirst, directly control instrumental performance by activating prepotent responses. In contrast, incentive-learning theory, originally espoused by Tolman (1949a, 1949b), maintains that the motivational effect of such states is indirect in that they act by determining the *incentive value* that an animal assigns to the instrumental reinforcer or outcome. Moreover, Tolman argued that this assignment is not automatic; rather, an animal has to learn about the value of a particular outcome in a given motivational state by experiencing it in that state. Dickinson and Balleine (1994) refer to this process as *incentive learning*.

Evidence for the role of incentive learning in the motivational control of instrumental action comes from a va-

riety of sources. Balleine (1992), for example, investigated the effect of devaluing a food outcome by shifts in

the level of food deprivation. Having trained hungry rats

to press a lever for food pellets, he compared performance in extinction following deprivation of the food supply in the home cage with that observed in the nondeprived state. Contrary to the prediction of drive theory, the level of deprivation on test had no detectable effect on performance during the extinction test; the undeprived rats pressed just as frequently as the hungry ones, indicating that, at least within the range of deprivation levels and with the training procedure employed by Balleine, the current motivational state of the animal has no direct effect on instrumental performance. In order to produce motivational control over test performance, Balleine had to give the rats prior exposure to the food pellets in the nondeprived state. Thus, only animals that were preexposed to the pellets in separate feeding cages in the nondeprived state prior to instrumental training showed a reduction of instrumental performance when tested in the nondeprived state. Balleine and Dickinson (1994) have recently replicated this incentive learning effect, using a procedure in which the animals were re- rather than preexposed to the outcome in that the incentive learning treatment was given between training and testing.

This research was supported by a grant from the SERC to A.D., and a Research Fellowship from Jesus College, Cambridge, to B.B. Correspondence should be addressed to A. Dickinson, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB4 3EB, England.

⁻Accepted by previous editor, Vincent M. LoLordo

This dependency of motivational control upon exposure to the outcome in the test state provides direct evidence for the role of incentive learning in the processes governing instrumental action. Such exposure allows animals to learn about the relatively low value of the outcome in the nondeprived state, learning which is reflected in their instrumental performance when they are subsequently tested in that state. And, in the absence of the opportunity for incentive learning, the shift to the low motivational state should have no detectable effect, for under these circumstances performance in the extinction test should reflect the high incentive value assigned to the outcome when the animals are hungry during instrumental training.

It is important to note that the incentive learning treatment could not have acted by altering the impact of the outcome itself in these studies, because performance was assessed in the absence of outcome presentations (i.e., in extinction). Rather, the incentive learning treatment must have operated by allowing the current motivational state to determine incentive value assigned to the outcome through its encoding in an animal's knowledge of the instrumental contingency. It is for this reason that incentive learning must be viewed as operating in concert with an account of instrumental learning, such as that espoused by Tolman (e.g., 1959), which assumes that conditioning reflects knowledge of the action—outcome relation.

Evidence favoring this interpretation comes from demonstrations that incentive learning is specific to the outcome associated with the instrumental action. In further studies, both Balleine (1992) and Balleine and Dickinson (1994) trained hungry rats to perform two different actions for different outcomes before testing performance in extinction in the absence of any deprivation. In all cases the animals showed a relative reduction in the performance of the action trained with the outcome that they had previously experienced in the nondeprived state, a result which demonstrates that incentive learning is mediated by the instrumental contingency. Without going into details (see Dickinson & Balleine, 1994, for a review), the same conclusion arises from studies that involve an increase in the level of food deprivation (Balleine, 1992; Balleine, Ball, & Dickinson, 1994), a decrement in water deprivation (Lopez, Balleine, & Dickinson, 1992), a shift between rather than within motivational states (Dickinson & Dawson, 1988, 1989), and finally, the induction of drug states (Balleine et al., 1994).

In summary, the studies of incentive learning indicate that primary motivational states, such as hunger and thirst, do not always have the direct impact on instrumental performance that classic drive theory envisages; rather, such states can control instrumental performance by determining the incentive value assigned to the outcome. This conclusion has been established, however, only for conditions in which the animals receive a limited amount of training (e.g., Balleine, 1992; Lopez et al., 1992), and there are reasons to believe that the role

of incentive learning may vary with the type of instrumental training. In a series of studies, Adams (1982) devalued the outcome following instrumental training by conditioning an aversion to it, in order to assess the effect of a variety of different instrumental training regimes on the devaluation effect. Although the overall pattern of results was complex, one generalization to emerge was that giving extended instrumental training distributed over a number of sessions reduces the sensitivity to devaluation. Since a motivational shift from training in a deprived state to testing in an undeprived state, when coupled with an incentive learning treatment, also serves to devalue the outcome, the effect of this treatment may also vary with the amount of instrumental training. The first of our studies addressed this issue directly.

EXPERIMENT 1

There were four stages in this study: instrumental training, reexposure, testing, and reacquisition (see Table 1). Initially all animals were trained to press a lever for food pellets while hungry in one of two training conditions. The limited training, or 120-outcome, condition followed that employed by Balleine (1992—Experiment 3A) and Balleine and Dickinson (1994—Experiment 1A) in that only four sessions of instrumental training were given, in each of which the animals were allowed to earn 30 pellets. By contrast, animals in the extended training, or 360-outcome, condition received a further 8 sessions of training before the incentive learning stage.

Whereas Balleine (1992) administered the incentive learning treatment prior to instrumental training, we followed the procedure employed by Balleine and Dickinson (1994) by giving this treatment between the end of instrumental training and test, thereby equating the time between incentive learning and testing in the two training conditions. During this reexposure stage, the animals had access to the food pellets in separate feeding cages when they were in either a relatively *high* or a relatively *low* state of hunger induced by varying the prior availability of food in their home cages. Thus, unlike the animals in the high condition, those in the low condition had the opportunity to learn about the relatively low value of the food pellets in the nondeprived state.

On the basis of the Balleine and Dickinson (1994) study we expected the effect of this reexposure to be

Table 1
Design of Experiment 1

Group	Training $H: lp \rightarrow Pel$	Reexposure		Test	Reacquisition
Low/low High/low		L:Pel L:0	H:0 H:Pel	$L: lp \to 0$ $L: lp \to 0$	$\begin{array}{c} L:lp \rightarrow Pel \\ L:lp \rightarrow Pel \end{array}$
Low/high High/high		L:Pel L:0	H:0 H:Pel	$\begin{array}{l} H:lp \rightarrow 0 \\ H:lp \rightarrow 0 \end{array}$	$\begin{array}{l} \text{H:lp} \rightarrow \text{Pel} \\ \text{H:lp} \rightarrow \text{Pel} \end{array}$

Note—H, high deprivation; L, low deprivation; lp, leverpress; Pel, pellet outcome; 0, no outcome.

manifest in the subsequent test when the animals were once again given the opportunity to press the lever, but this time in extinction. Specifically, in the 120 condition the animals that received both reexposure and testing in the low state were expected to press less than the animals tested following food deprivation. Thus, if the first term in the group designations refers to the motivational state during reexposure and the second term to the state during testing, our predictions were that Group Low/Low should press less on test than Groups Low/High and High/ High, which in turn should not differ as a function of the reexposure treatment. By contrast, Group High/Low, which received reexposure in the deprived state but was tested in the nondeprived state, should press just as much as Groups Low/High and High/High if outcome devaluation by motivational shift depends upon incentive learning. Given the replication of this pattern of results, the question of prime interest was whether a similar devaluation effect would be observed after more extended instrumental training in the 360 condition. On the basis of Adams's (1982) failure to detect an effect of aversion conditioning to the outcome after extended training, we anticipated that all animals in the 360 condition would press at equivalent, high rates on test.

Finally, all animals were given a reacquisition session under the test state to confirm that the different levels of deprivation did have an impact on the incentive value of the food pellets as indexed by their reinforcing property.

Method

Subjects and Apparatus. Thirty-two naive adult male hooded Lister rats of at least 90 days of age and 250 gm in weight were housed in pairs with unlimited access to water in their home cages. Instrumental training and testing took place in four Campden Instrument operant chambers, each equipped with a dispenser that could deliver 45-mg Noyes pellets (improved Formula A) into a recessed magazine that the rats could enter through a flap door positioned in the center of the front wall. A retractable lever was located to the right of the flap door, and each chamber was illuminated by a 3-W 24-V houselight mounted on the front wall above the flap door. A BBC microcomputer, equipped with the Spider extension for on-line control (Paul Fray Ltd, England), controlled the equipment and recorded the leverpresses. Reexposure took place in four black plastic cages, 28 × 12 × 12 cm, each of which contained a glass dish in which the food pellets were placed.

Procedure. Food was removed from the home cages on the day prior to the start of training, and all subjects were given access to their maintenance diet in their home cages for 1.5 h following each daily training session. Half the animals were assigned to the 120 condition and the remainder to the 360 condition. The latter animals received two 30-min sessions of magazine training in the operant chambers with the lever retracted on Days 1 and 2. During each of these sessions, pellets were delivered on a random time (RT) 60-sec schedule. There followed 12 sessions of instrumental training, each of which started with the insertion of the lever and the onset of the houselight and terminated, after 30 pellets had been delivered, with the retraction of the lever and the offset of the houselight. Leverpressing was reinforced on a random interval (RI) schedule whose parameter increased from 2 sec through 15 sec to 30 sec across the first three sessions (Days 3-5). The animals in the 360 condition then received 9 further sessions of training (Days 6-14) on the RI 30-sec schedule.

During Days 1–3 the animals in the 120 condition remained in their home cages. Over the next 5 days, however, they were treated in a way designed to match approximately the handling and transport conditions of the animals in the 360 condition; each day they were removed from their home cages and placed individually into similar cages but in the adjacent colony room, before being returned to their home cages 1–2 min later. Instrumental training began on Day 9 and was identical to that received by animals in the 360 condition, except that only two sessions of training on the RI 30-sec schedule were given in the 120 condition. In addition, to familiarize the animals with the feeding cages prior to incentive learning, immediately following the instrumental training session on Day 10 they all received a single 15-min exposure to the feeding cages, during which they were allowed to eat 30 pellets.

Across the 4 days following instrumental training (Days 15-18), the animals were given the reexposure treatment during which their deprivation state was alternated by giving them free access to home-cage food on the nondeprived days and removing the food on the deprived days. This deprivation schedule was initiated immediately after the 1.5-h access period that followed the last instrumental training session on Day 14, and the order of the deprivation states was counterbalanced within each training condition. The reversal of the deprivation state occurred when the animals were returned to their cages after being placed in the feeding cages for 15 min each day. Half the animals in each trainingcondition/deprivation-order subgroup were given 30 pellets in the feeding cages on days when they were deprived and none on days when they were underrived (Groups High/Low and High/High). By contrast, the remaining animals received access to the food pellets on the days when they were undeprived (Groups Low/Low and Low/High).

Following the last placement in the feeding cages on Day 18, half the animals in each counterbalanced condition were given free access to home-cage food for the rest of the experiment (Groups Low/Low and High/Low), whereas the remaining animals were returned to the 1.5-h daily feeding schedule (Groups Low/High and High/High). Two 30-min instrumental test sessions were then given; the test on Day 19 was conducted in extinction, whereas leverpressing was reinforced on the RI 30-sec schedule on Day 20. The number of leverpresses performed in each 5-min period of these tests was recorded.

Statistical Analysis. The number of leverpresses during the tests was analysed in a four-way, mixed analysis of variance. The three between-subjects factors distinguished between animals in the 120 and 360 training conditions (training), between those that received reexposure to the pellets in the high and low states (reexposure), and between those that were tested in the high and low states (test). The within-subjects factor (bin) distinguished between performance in successive 5-min bins of the test sessions. Because the variance in the rates of leverpressing decreased with the decline in the overall rate during the extinction test, this measure was subject to a square-root transform prior to all analyses in both this and the subsequent experiments. The mean square-root leverpresses per minute are reported in parentheses where appropriate. The reliability of the observed effects was evaluated against the conventional criterion of p = .05.

Results and Discussion

The results of prime interest, presented in Figure 1, are the number of leverpresses performed by the various groups during the extinction test. The performance of the animals in the 120 condition replicated the incentive learning effect observed by Balleine and Dickinson (1994) in that, at least during the initial period of the session,

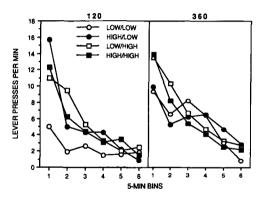


Figure 1. Experiment 1: mean number of leverpresses per minute during the extinction test following training in the 120 condition (left panel) and the 360 condition (right panel). The animals were tested in either the low (Groups Low/Low and High/Low) or the high (Groups Low/High and High/High) deprivation state following resposure to the food pellets in either the low (Groups Low/Low and Low/High) or the high (Groups High/Low and High/High) deprivation state.

the animals in Group Low/Low pressed at a lower rate than did the other groups, which did not differ among themselves. Thus, a reduction in the relevant motivational state only attenuated instrumental performance after limited training if the animals had received prior exposure to the outcome in that state. By contrast, a very different pattern of results was observed after more extended training in the 360 condition. In this condition, the main variable affecting performance appeared to be the motivational state on test, with animals tested in the low state, Groups Low/Low and High/Low, pressing less at the outset of the extinction session than those tested in the high state, Groups Low/High and High/High.

This description was confirmed by statistical analysis. An overall analysis yielded a significant test \times bin interaction [F(5,120) = 2.68] and a marginal four-way interaction [F(5,120) = 2.04, p < .08]. These interactions prompted separate three-way analyses for each 5-min bin. The analysis for the first bin produced a significant training \times reexposure \times test interaction [F(1,24) =4.47], which in turn led to a consideration of the effects of the reexposure and the test states in each training conditioning. These two factors interacted significantly in the 120 condition [F(1,24) = 8.61], and Newman-Keuls tests showed that Group Low/Low (2.21) pressed less in the first 5 min of the extinction test than the remaining groups, Groups High/Low (3.95), Low/High (3.28), and High/High (3.47), which in turn did not differ. By contrast, the analysis of performance during the first 5 min of extinction in the 360 training condition yielded only a significant effect of the test state [F(1,24) = 4.75], with animals tested in the high state (3.76) pressing more than those tested in the low state (3.07). The only significant effects in analyses of performance during the remaining periods was one of training condition in Bins 3 and 4 [Fs(1,24) > 4.77], reflecting the fact that the animals in

the 360 condition (2.36) showed greater resistance to extinction than did those in the 120 condition (1.77).

The effects observed in the extinction test were not compromised by differences in performance established during instrumental training. On the last day of training, the means of the number of leverpresses per minute in the 120 condition were 14.6 (3.79), 12.3 (3.47), 14.5 (3.78), and 12.7 (3.56) for Groups Low/Low, High/Low, Low/High, and High/High, respectively. The comparable rates in the 360 condition were no higher, being 15.8 (3.93), 12.9 (3.58), 14.2 (3.73), and 13.8 (3.67) presses per minute, respectively. An overall analysis employing the same three factors as those used with the extinction data revealed no significant main effects or interactions [largest F(1,24) = 1.67].

Finally, performance during the reacquisition test verified that the difference in the high- and low-deprivation conditions modulated the reinforcing properties of the food pellets in that the animals tested under the high state pressed more than those tested under the low state [F(1,24) = 18.75]. In addition, higher rates of pressing were observed after extended training in the 360 condition [F(1,24) = 6.43]. No effect of the reexposure treatment was detected [F(1,24) = 1.39], however, nor were there any significant interactions [largest F(1,24) =1.01]. Balleine (1992) and Balleine and Dickinson (1994) also failed to detect any significant effect of the incentive learning treatment during reacquisition, presumably reflecting the fact that the animals very rapidly readjust the incentive value assigned to the outcome on the basis of their experience with the outcome during the early part of the session. The mean rates during the reacquisition session in the 120 condition were 7.2 (2.64) and 5.0 (2.17) presses per minute for animals retrained in the high (Groups Low/High and High/High) and low states (Groups Low/Low and High/Low), respectively, and 9.8 (3.07) and 6.1 (2.42) presses per minute for the corresponding groups in the 360 condition.

In summary, this study has established two main findings. The first is the replication of the prior demonstrations (e.g., Balleine, 1992; Balleine & Dickinson, 1994) that motivational control of instrumental performance depends on incentive learning. A decrement in an animal's level of hunger produces a consequent reduction in instrumental performance after limited training only if the animal has received prior experience of the outcome in the test state. The major finding, however, concerns the effect of extended training on the susceptibility of instrumental performance to outcome devaluation. Just as Adams (1982) demonstrated that extended training reduces sensitivity to outcome devaluation based on aversion conditioning, we have found a comparable reduction in the case of devaluation through a motivationally based incentive learning treatment. After extended training, the incentive learning treatment had no detectable effect on leverpressing during the test; rather, to the extent that motivational variables had any impact on instrumental performance, the critical factor was the current motivational state, with animals pressing somewhat faster following deprivation.

EXPERIMENT 2

Adams (1982) went some way toward determining the factors underlying the variation in sensitivity to outcome devaluation. As he pointed out, extended training confounds potentially important variables: the number of instrumental actions performed; the number of reinforced actions; and the number of outcomes presented. One of his experiments (Experiment 5) identified the importance of the last of these factors. The critical condition in this study matched the total number of outcome presentations to that received by animals in the extended training condition, while equating the instrumental training to that experienced by animals in the limited training conditioning. This was achieved by giving an appropriate number of sessions in the operant chamber during which the outcome was presented noncontingently in the absence of the lever prior to limited instrumental training. While replicating the basic devaluation effect in a group given solely limited instrumental training, Adams failed to observe any such effect in the animal given noncontingent exposure prior to this instrumental training.

The purpose of our second study was to investigate whether a comparable insensitivity could be engendered by noncontingent training when the outcome was devalued by the incentive learning treatment combined with a motivational shift. To this end, two groups of animals received either the 120 or the 360 instrumental training conditions employed in the first study prior to the reexposure treatment in either the high (Group High) or the low state (Group Low). On the basis of our prior results, we expected Group Low to press significantly less than Group High when tested in the low state in the 120 condition, but not in the 360 condition. The animals in the novel, 240/120 condition received the same instrumental training as did those in the 120 condition; prior to this training, however, these animals received 240 noncontingent presentations of the outcome in the operant chambers so that the total number of outcomes received matched that in the 360 condition. The critical issue is whether, following training in this 240/120 condition and the reexposure treatment, Group Low, the animals that had received reexposure in the low state. would press less than Group High when tested in the nondeprived state. A generalization from the resistance to outcome devaluation observed by Adams (1982) following a similar training regime anticipates little evidence for an incentive learning effect in the 240/120 condition.

Method

Subjects and Apparatus. Twenty-four naive male hooded Lister rats with age and weight comparable to those of rats in Experiment 1 were housed under the same conditions and trained in the same apparatus.

Procedure. The animals were allocated equally to three different training conditions. Those in the 120 and 360 conditions were trained with the same procedure as that employed for the corresponding condition in Experiment 1. That is, after two sessions of magazine training and two of instrumental acquisition, the animals received further instrumental training sessions in which leverpressing was reinforced with food pellets on an RI 30-sec schedule. The training was scheduled so that the four sessions of instrumental training in the 120 condition terminated on the same day (Day 14) as did the 12 sessions in the 360 condition. Training in the third, 240/120 condition started at the same time as that in the 360 condition with two sessions of magazine training. Thereafter, however, these animals received 8 sessions in each of which outcomes were delivered on a RT 30-sec schedule with the lever retracted followed by four sessions of instrumental training. During these sessions, leverpressing was reinforced on an RI schedule whose parameter value was 2 sec on the first session, 15 sec on the second, and 30 sec on the last two, so that the instrumental training regime was identical to that employed in the 120 condition. Each session terminated after 30 outcomes had been delivered. All animals were maintained on a food deprivation schedule throughout training by being fed for 1.5 h in the home cage after every session.

The reexposure treatment followed the procedure employed in Experiment 1 except in one respect: the rats remained in their home cages on the days when they did not receive exposure to the food pellets in the feeding cages. This exposure occurred on days when the animals had been deprived of food in their home cages during the preceding 22.5 h in Group High and on days when they were not so deprived in Group Low, with the order of the deprivation states counterbalanced within each training condition. Following the last day of reexposure, all animals were given free access to their maintenance diet before receiving a 30-min extinction test on the next day.

Results and Discussion

Figure 2 illustrates the results of prime interest: the rates of leverpressing during the extinction test. These are given separately for each of the three training conditions. As in the first study, the main effect of the reexposure treatment was evident in the first half of the test session, where there was a clear replication of the incentive learning effect in the 120 condition; Group Low

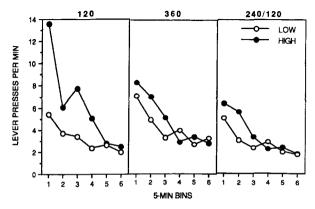


Figure 2. Experiment 2: mean number of leverpresses per minute during the extinction test following training in the 120 condition (left panel), the 360 condition (middle panel), and the 240/120 condition (right panel). The animals were tested in either the low or the high deprivation state.

pressed less than Group High. Although the difference between these groups was in the same direction in the 360 condition, it was neither reliable nor as large as that in the 120 condition (see below). The novel, 240/120 condition produced a pattern of extinction performance very similar to that observed in the 360 condition, with no reliable evidence that Group Low pressed consistently less than Group High.

The overall analysis of extinction performance established the reliability of the effect of the incentive learning treatment [F(1,42) = 7.78], an effect that varied across the extinction test [F(5,210) = 3.50]. Separate analyses of performance during each 5-min bin of the test produced a significant interaction between the training condition and the reexposure treatment during the first bin [F(2,42) = 4.63]. Simple main effects analyses within each training condition showed that Group High (3.58) pressed significantly more than Group Low (2.29) in the 120 condition during the first 5 min [F(1,42) = 19.99], but not in the 360 and 240/120 conditions (F < 1) in both cases).

The analyses of performance during the remaining bins confirmed the pattern observed during the first 5 min of the extinction test. There was a significant main effect of reexposure during Bins 2 and 3 [Fs(1,42) > 5.67], and a significant training × reexposure interaction [F(2,42) = 3.48] during Bin 4. Simple main effect analyses revealed that the effect of reexposure was significant in the 120 condition during Bins 3 (High, 2.75; Low, 1.65) and 4 [High, 2.08; Low, 1.42; Fs(1,42) > 5.05]. By contrast, there were no significant effects of reexposure in the 360 and 240/120 training conditions throughout the extinction test [Fs(1,42) < 2.87].

In general, this pattern of results corresponds to that observed by Adams (1982) in the case of outcome devaluation brought about by aversion conditioning rather than a combination of motivational shift and incentive learning. A reliable devaluation effect is observed after limited exposure to the outcome during training, but not after more extended exposure whether or not the outcome is contingent on an instrumental action. This difference cannot be traced to variations in performance at the end of training; the effects of the training condition [F(2,42) = 2.08] and of the reexposure treatment (F < 1)on the rates of pressing on the last training session were not significant, nor was their interaction (F < 1). The means of the number of leverpresses per minute in this session were, for Groups High and Low, respectively, in the 120 condition 11.6 (3.31) and 11.1 (3.28); in the 360 condition 14.7 (3.74) and 16.1 (3.98); and in the 240/120 condition 13.3 (3.55) and 13.4 (3.61).

EXPERIMENT 3

Although the previous two experiments failed to detect an effect of the incentive learning treatment after extended training, what is less clear is whether performance after such training is completely impervious to outcome devaluation. In this respect, it is notable that in both the

extended training conditions of Experiment 2 (but not in Experiment 1) the animals reexposed to the food pellets in the low state tended to press less on test than those deprived during reexposure, although not significantly so. This trend led us to reexamine the status of instrumental performance after overtraining by investigating the effect of giving more extensive reexposure treatment. To this end, the total number of food pellets presented during reexposure was increased from 60 to 200.

In addition, a further modification was made to the reexposure procedure in an attempt to enhance its impact. In the previous two studies, the reexposure was administered in separate feeding cages on the assumption that learning about the relatively low value of the pellets in the nondeprived state would transfer from the feeding cages to the operant chambers. There is reason to believe, however, that such transfer may be far from perfect. Capaldi and Myers (1978), for example, observed that undergived rats readily at reward pellets in the goalbox of a runway, in which they had received extensive prior training under hunger, even though they had been sated on the same pellets in their home cages immediately prior to placement in the runway. Capaldi and Meyers (1978) interpreted this finding as evidence that feeding comes under the control of contextual stimuli. If the same is true of the incentive value assigned to a particular food, we should expect to observe a stronger incentive learning effect if exposure to the pellets in the low state takes place in the operant chambers rather than in the feeding cages. To determine the potential effect of this factor, half the animals received reexposure in the operant chambers rather than in the feeding cages.

In summary, all animals received extended instrumental training (360 condition) prior to the reexposure treatment in either the feeding cages or the operant chambers and a subsequent extinction test in the non-deprived state. A reliable incentive learning effect would be manifest if the animals reexposed to the food in the nondeprived state, Group Low, pressed significantly less on test than those that were reexposed to the food when hungry, Group High.

Method

Subjects and Apparatus. Thirty-two adult male hooded Lister rats were trained and tested under the same housing conditions and in the same apparatus as those employed in the previous studies.

Procedure. The instrumental training procedure followed exactly that employed in the 360 conditions of the previous studies. Following two magazine training sessions and two instrumental acquisition sessions, all animals received 10 sessions in which leverpressing was reinforced on an RI 30-sec schedule for 30 outcomes per session. Throughout this period, the rats were maintained on a 22.5-h food deprivation schedule.

The incentive learning stage was also similar to that employed previously in that all animals were subject to a counterbalanced alternation between food deprivation and nondeprivation across 4 days, with half the animals receiving access to the food pellets on days when they were undeprived, Group Low, and the remainder on days when they were hungry, Group High. The procedure was modified in two respects from that used previously, however.

Whereas half of the animals in each group were exposed to the pellets in the feeding cages as in the prior studies, the rest were placed in the operant chambers, where the pellets were delivered noncontingently on an RT 30-sec schedule. Moreover, each animal received two placements in either the feeding cages or the operant chambers on each of the days on which they were scheduled to receive access to the food pellets, one in the morning and one in the afternoon. Fifty pellets were presented during each of these placements.

Finally, all animals were given a 30-min instrumental extinction test in the nondeprived state.

Results and Discussion

As Figure 3 shows, the results of this study were straightforward. The overall analysis vielded a significant interaction between the reexposure treatment and the 5-min bins of the extinction test [F(5,140) = 5.09], due to the fact that Group High (4.91) pressed more than Group Low (3.05) during the third 5-min bin [F(1.28) = 18.87]. There was no reliable effect of the reexposure treatment during any other period of the extinction test [Fs (1.28) < 3.76]. Thus animals that had received prior exposure to the food pellets in the nondeprived state, Group Low, pressed less than those that had received the same exposure when hungry. Simple main effects analyses established that the effect of the reexposure treatment was reliable when administered both in the feeding cages [High, 4.87; Low, 3.10; F(1.28)= 6.41], and in the operant chambers [High, 4.85; Low, 3.00; F(1,28) = 4.39].

There was no evidence, however, that the magnitude of the reexposure effect depended on the location of the incentive learning treatment; the F ratios for the interactions involving the reexposure location and the reexposure treatments were both less than one. Thus, it would appear that incentive learning generalizes fully from the feeding cages to the operant chambers. This finding accords with the generalization of incentive learning observed by Balleine and Dickinson (1991) in the case of outcome devaluation by aversion conditioning. They reported that establishing an aversion to an out-

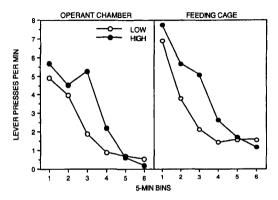


Figure 3. Experiment 3: mean number of leverpresses per minute during the extinction test following reexposure to the food pellets in either the operant chambers (left panel) or the feeding cages (right panel). The animals were tested in either the low or the high deprivation state.

come by following a single session of instrumental training with an isotonic lithium chloride injection only produces a devaluation effect in a subsequent extinction test if the animal receives prior exposure to the outcome following the injection. The magnitude of this devaluation effect was unaffected, however, by whether the exposure occurred in the operant chambers or in the same feeding cages as those employed in our study. In fact, the only detectable effect of the location of reexposure in the present study was that animals that received this treatment in the operant chambers (3.07) pressed somewhat less on test than those that received reexposure in the feeding cages [3.64; F(1,28) = 5.32].

An analysis of performance on the last training sessions yielded F ratios for the main effects and interaction involving reexposure location and treatment that were all less than one. The mean presses per minute for Groups High and Low were 11.6 (3.39) and 11.8 (3.41), respectively, when reexposure occurred in the operant chamber, and 13.9 (3.72) and 12.4 (3.43), respectively, for rats reexposed in the feeding cages.

In conclusion, this study has established that a reliable outcome devaluation effect can be detected after extended instrumental training following more extensive incentive learning treatment.

GENERAL DISCUSSION

A number of conclusions about the motivational control of instrumental performance can be drawn from these studies. The first concerns the role of incentive learning after limited training. Experiment 1 replicated our previous finding that, following training under food deprivation, the motivational state on test has no direct impact on instrumental performance (Balleine, 1992; Balleine & Dickinson, 1994). Undeprived animals press just as frequently as hungry ones in the absence of prior experience with the food outcome in the nondeprived state. In order for the current motivational state to exert control over instrumental performance, prior experience of the outcome in this state is required. This finding accords with the predictions of incentive-learning theory.

By contrast, the pattern of test performance observed after extended training provides some evidence for a direct impact of the current motivational state; the high level of deprivation produced more leverpresses on test than did the low one, although the effect was small in magnitude. More important, however, is the fact that extended training decreases the impact of the incentive learning treatment. In both Experiments 1 and 2, the amount of reexposure to the food pellets in the low state that produced a reduction in pressing on test after limited training had no reliable effect on performance after more extended training. This is not to claim, however, that overtraining renders performance entirely resistant to outcome devaluation, for a reliable incentive learning effect was observed in the final study.

Perhaps the most straightforward explanation of this pattern of results can be derived from the analysis of resistance to satiation offered by Capaldi, Davidson, and Meyers (1981: see also Morgan, 1974). They exposed hungry rats to reward pellets in their home cages prior to differential training with this reward in a runway in the underrived state. In contrast to rats that had received the preexposure in a sated state, these animals ran consistently faster on rewarded trials than on nonrewarded ones, a finding that Capaldi et al. (1981) interpret as evidence that prior exposure to the pellets when animals are hungry enhances their reinforcing property when they are sated. To the extent that runway performance reflected the value assigned by the animals to the food outcome, this study suggests that the high value assigned to the food pellets when the animals are hungry generalizes to the assignment made when the animals are undeprived. Indeed, we have to assume just such generalization if we are to explain why the animals in the 120 condition of Experiment 1 which were reexposed in the deprived state but tested in the nondeprived state, Group High/Low, pressed just as much on test as those tested following deprivation, Groups Low/High and High/ High, Furthermore, if it is assumed that the strength of this generalization increases with the amount of exposure to the food under hunger, it is to be expected that a given incentive learning treatment should be less effective in reducing its value.

Two features of the present results clearly accord with this account. The first is that the critical determinant of the overtraining effect is the number of exposures to the pellets when the animals are hungry during training rather than amount of instrumental conditioning (Experiment 2). Second, an incentive learning effect can be detected after extended training if the amount of reexposure given animals in the nondeprived state is increased (Experiment 3). Presumably, if the incentive learning treatment is to result in the assignment of a low value, greater reexposure is required to counteract the enhanced generalization of value from the hungry state after extended training. It should be noted, however, that the reliability of the reexposure effect observed in Experiment 3 may well reflect the greater statistical power of this study rather than the increased reexposure received by the animals.

One problem for this generalization account, however, arises from performance of the animals trained and reexposed in the deprived state but tested in the nondeprived state. If more extended exposure to pellets under the high state enhances the generalization of incentive value to the low state, Group High/Low in Experiment 1 and Group High in Experiment 2 should have pressed more on test after extended training than after limited training. The test performance in these groups should have reflected directly the generalization of incentive value assigned during training (and reexposure) in the high state to the control of performance in the low state. As Figures 1 and 2 show, however, just the reverse was the case; a post hoc analysis of the combined performance of these two groups revealed that animals in the 120 condition pressed at a mean rate of 14.3 (3.70)

per minute during the first 5 min of extinction, which was significantly higher than the mean rate of 8.9 (2.92) presses per minute observed in the 360 condition [F(1,22) = 7.88]. A related difficulty arises from the emergence of a direct effect of motivational state on extinction performance after extended training in Experiment 1. Indeed, to the extent that extended training enhances the generalization of a high incentive value from the deprived to the undeprived state, one might expect the current motivational condition to have relatively less impact after such training.

The variation in the direct impact of a motivational shift with the amount of training suggests that the predominant process controlling performance may well change as a function of the amount (and nature) of the training. Although this is a venerable idea dating back at least to James's (1890) classic analysis of habits (see Boakes, 1993), Dickinson (1985, 1989) offered a contemporary statement specifically addressed to the variations in the sensitivity to outcome devaluation. Dickinson argued that instrumental training engages two independent learning processes which operate concurrently. The first is a simple habit process that follows the general principles of classic stimulus-response $(S \rightarrow R)$ reinforcement theory. As each presentation of a contiguous reinforcer serves to strengthen the propensity to perform its associated response, the contribution of this $S \rightarrow R$ process to instrumental performance should increase with the amount of training, as is illustrated in Figure 4.

The second process is based on knowledge of the action—outcome $(A \rightarrow O)$ relation. In order to account for the influence of a variety of training regimes on the devaluation effect, Dickinson (1985, 1989) suggested that this knowledge represents the *experienced* correlation between the rate of responding and the rate of reinforcement, a variable that Baum (1973) had previously

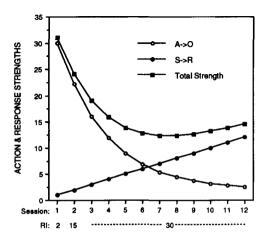


Figure 4. Dual-process theory of overtraining: changes in the contributions of the $A \rightarrow O$ process and the $S \rightarrow R$ process to the total strength of instrumental behavior, as a function of the number of sessions of training on the schedules employed in the present experiments (see text for further explanation).

identified as critical in instrumental conditioning. Given this characterization of action-outcome learning, the contribution of this process to instrumental performance should vary during the type of training given in the present study. For the sake of illustration, we have assumed that on the interval schedules employed in the present studies the contribution of the A \rightarrow O process is determined by the programmed rate of reinforcement, so that the strength of this process at the end of the first training session on the RI 2-sec schedule is 30 (see Figure 4). At the end of subsequent sessions, the strength of this process is assumed to be a weighted average of the outcome rate (O) in that session and the $A \rightarrow O$ strength at the end of the previous session, so that the strength of the process at the end of session n, $(A \rightarrow O)_n$, is given by: $a \cdot O_n + (1-a) \cdot (A \rightarrow O)_{n-1}$. Figure 4 illustrates the decreasing strength of the A \rightarrow 0 process across the extended training regime given in the present studies when a = 0.3. According to this dual-process theory, the overall rate of instrumental performance at any given stage of training will be determined by the total strength produced by summing the strengths of the $A \rightarrow O$ and $S \rightarrow R$ processes.

Changes in the relative contributions of these two processes with training have straightforward consequences for variations in the sensitivity to outcome devaluation. Given that outcome devaluation can only operate through knowledge of the action—outcome relation, devaluation should be effective to the extent that the $A \rightarrow O$ process makes a significant contribution to the total strength. As Figure 4 shows, this contribution is much greater after the four sessions of training in the 120 condition than after the 12 sessions in the 360 condition, with the consequence that an outcome devaluation effect should be more difficult to detect after extended training, as we observed in the first two studies.

The dual-process theory also provides an account of variations in the sensitivity of performance to the direct impact of the animal's current motivational state. Drive theory (e.g., Hull, 1943) and accounts of the motivational influence of Pavlovian processes on instrumental performance (e.g., Rescorla & Solomon, 1967) have traditionally assumed that motivational states exert their direct effects through an $S \rightarrow R$ process, and thus we should expect the impact of such states to increase with extended training, the effect observed in Experiment 1. It is this fact that explains why a reduction in deprivation to the low state in Experiments 1 and 2 (in the absence of reexposure to the outcome in the low state) had a greater impact on performance after extended training. Although the total strength may be comparable after limited (4 sessions) and extended (12 sessions) training, the $S \rightarrow R$ process makes a much greater contribution in the extended condition, with the consequence that a variable mediated by this process should exert a more substantial effect with training.

Finally, this account also anticipates the resistance to devaluation engendered by prior noncontingent presentations of the outcome observed in Experiment 2. As has been suggested previously in the context of learned helplessness (see, e.g., Maier & Seligman, 1976), prior noncontingent presentation of an instrumental outcome should interfere with subsequent learning of an action-outcome contingency or correlation. Thus, the prior noncontingent training given to Group 240/120 should have interfered with the development of the $A \rightarrow O$ process during instrumental training so that in this training condition performance should have been primarily based on the $S \rightarrow R$ process rather than the $A \rightarrow O$ process with the consequent insensitivity to outcome devaluation.

This dual-process theory can also offer an account of an apparent discrepancy in the reported effects of overtraining on sensitivity to devaluation by aversion conditioning. In contrast to Adams's (1982) findings, Colwill and Rescorla (1985) failed to detect any effect of extended training on resistance to outcome devaluation: indeed, if anything, sensitivity to devaluation increased with training in their procedure (Colwill & Rescorla, 1988). A notable feature of this procedure, however, is that the animals always received training with two different outcomes contingent upon different actions. Thus, for example, in their first study (Colwill & Rescorla, 1985), action A_m received moderate training and action A_e extended training, both with the same outcome, O₁. The subsequent conditioning of an aversion to O₁ produced comparable decrements in the performance of A_m and A_e in an extinction test. If it is assumed, however, that action-outcome knowledge is based on the experienced correlation between action and outcome rates, this result is not surprising, because, interspersed with the training of A_m and A_e, a third, control action, A_c, was reinforced with a different outcome, O2, in separate training sessions. Consequently, the animals were exposed to a relatively strong positive correlation between the rates of both A_m and A_e and the rate of O₁ throughout training—O₁ occurred in sessions when these actions were performed, but never in sessions when only the manipulandum for A_c was available. Under this training regime. correlational theory would predict that knowledge of the A_e-O₁ relation should be sustained however extended the training, thus rendering this action permanently sensitive to outcome devaluation.

Whatever the merits of this account, the present studies do show that the motivational control of instrumental performance is complex and depends on the training regime. After limited training, the current motivational state of the animal has little direct control over performance, which depends on prior experience of the outcome in the test state. By contrast, extended training on either contingent or noncontingent schedules reduces the role of incentive learning, while enhancing the direct impact of the current motivational state on performance.

REFERENCES

ADAMS, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, 34B, 77-98.

- BALLEINE, B. (1992). Instrumental performance following a shift in primary motivation depends upon incentive learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 236-250.
- BALLEINE, B., BALL, J., & DICKINSON, A. (1994). Benzodiazepine-in-duced outcome revaluation and the motivational control of instrumental action in rats. Behavioral Neuroscience, 108, 573-589.
- Balleine, B., & Dickinson, A. (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *Ouarterly Journal of Experimental Psychology*, **43B**, 279-296.
- BALLEINE, B., & DICKINSON, A. (1994). Role of cholecystokinin in the motivational control of instrumental action. *Behavioral Neuro-science*, 108, 590-605.
- BAUM, W. M. (1973). The correlational-based law of effect. Journal of the Experimental Analysis of Behavior, 20, 137-143.
- BOAKES, R. A. (1993). The role of repetition in transforming actions into habits: The contribution of John Watson and contemporary research to a persistent theme. *Mexican Journal of Behavior Analysis*, 19, 67-90.
- CAPALDI, E. D., DAVIDSON, T. L., & MYERS, D. E. (1981). Resistance to satiation: Reinforcing effects of food and eating under satiation. *Learning & Motivation*, 12, 171-195.
- CAPALDI, E. D., & MEYERS, D. E. (1978). Resistance to satiation of consummatory and instrumental performance. *Learning & Motiva*tion, 9, 197-201.
- COLWILL, R. M., & RESCORLA, R. A. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. Journal of Experimental Psychology: Animal Behavior Processes, 11, 520-536
- COLWILL, R. M., & RESCORLA, R. A. (1988). The role of responsereinforcer associations increases throughout extended instrumental training. *Animal Learning & Behavior*, 16, 105-111.
- DICKINSON, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Soci*ety of London: Series B, 308, 76-78.
- DICKINSON, A. (1989). Expectancy theory in animal conditioning. In S. B. Klein & R. R. Mowrer (Eds.), Contemporary learning theo-

- ries: Pavlovian conditioning and the status of traditional learning theories (pp. 279-308). Hillsdale, NJ: Erlbaum.
- DICKINSON, A., & BALLEINE, B. (1994). Motivational control of goal-directed action. Animal Learning & Behavior, 22, 1-18.
- DICKINSON, A., & DAWSON, G. R. (1988). Motivational control of instrumental performance: The role of prior experience of the reinforcer. Quarterly Journal of Experimental Psychology, 40B, 113-134
- DICKINSON, A., & DAWSON, G. R. (1989). Incentive learning and the motivational control of instrumental performance. *Quarterly Jour*nal of Experimental Psychology, 41B, 99-112.
- HULL, C. L. (1943). Principles of behavior. New York: Appleton-Century-Crofts.
- JAMES, W. (1890). The principles of psychology. New York: Holt.
- LOPEZ, M., BALLEINE, B., & DICKINSON, A. (1992). Incentive learning and the motivational control of instrumental performance by thirst. *Animal Learning & Behavior*, 20, 322-328.
- MAIER, S. F., & SELIGMAN, M. E. P. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, 105, 3-46.
- MORGAN, M. J. (1974). Resistance to satiation. Animal Behaviour, 22, 449-466.
- RESCORLA, R. A., & SOLOMON, R. L. (1967). Two-process learning theory: Relationship between Pavlovian conditioning and instrumental learning. *Psychological Review*, 74, 151-182.
- TOLMAN, E. C. (1949a). There is more than one kind of learning. *Psychological Review*, **56**, 144-155.
- TOLMAN, E. C. (1949b). The nature and function of wants. Psychological Review, 56, 357-369.
- TOLMAN, E. C. (1959). Principles of purposive behavior. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 92-157). New York: McGraw-Hill.

(Manuscript received March 20, 1994; revision accepted for publication June 21, 1994.)