

Motivational control of goal-directed action

ANTHONY DICKINSON and BERNARD BALLEINE
University of Cambridge, Cambridge, England

The control of goal-directed, instrumental actions by primary motivational states, such as hunger and thirst, is mediated by two processes. The first is engaged by the Pavlovian association between contextual or discriminative stimuli and the outcome or reinforcer presented during instrumental training. Such stimuli exert a motivational influence on instrumental performance that depends upon the relevance of the associated outcome to the current motivational state of the agent. Moreover, the motivational effects of these stimuli operate in the absence of prior experience with the outcome under the relevant motivational state. The second, instrumental, process is mediated by knowledge of the contingency between the action and its outcome and controls the value assigned to this outcome. In contrast to the Pavlovian process, motivational states do not influence the instrumental process directly; rather, the agent has to learn about the value of an outcome in a given motivational state by exposure to it while in that state. This incentive learning is similar in certain respects to the acquisition of "cathezes" envisaged by Tolman (1949a, 1949b).

The general adaptive significance of the capacity for goal-directed action is so obvious as to require little or no comment. It is this capacity that allows us and other animals to control our environment in the service of our desires and needs. And yet, if asked the most simple questions about this capacity, such as why does an animal perform an instrumental action for a food reward more readily when hungry rather than sated, a contemporary psychologist could tell us little more than Hull (1943) or Tolman (1949a, 1949b) nearly half a century ago. The fact is that the study of the psychological processes controlling the performance of simple, goal-directed, instrumental actions by basic primary motivational states has been neglected over the intervening decades.

It is true that we know much more about the neurophysiological and chemical mechanisms regulating consummatory behavior, such as eating, drinking, and copulating, and indeed about the major role of learning in such behaviors. A contemporary psychologist could also tell us much, at least on a functional level, about how instrumental behavior adapts to schedule and foraging constraints, and indeed about the effects of motivational variables on such adaptation. Moreover, there is an extensive literature on the psychological resources brought to bear in solving complex problems and decisions in the service of cognitive goals. But even so, we should search in vain among the literature for a consensus about the psychological processes by which primary motivational states, such as hunger and thirst, regulate simple goal-directed acts. This is the issue that we wish to address in the present paper.

This research was supported by the S.E.R.C. and the Commonwealth Scholarship Commission. We thank A. Watt and R. A. Boakes for their comments on a draft of the manuscript. Correspondence should be addressed to A. Dickinson, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.

GOAL-DIRECTED ACTION

A prerequisite for this discussion is a clear specification of what we mean by "goal directed." Our conception of a goal-directed action is psychological, being based upon the nature of the representations mediating performance. By characterizing an action as being "directed" at a goal, we mean that performance is mediated by knowledge of the contingency between the action and the goal or outcome, whether this knowledge is conceived of as an expectation or belief (e.g., Bolles, 1972; Dickinson, 1980; Irwin, 1971; Tolman, 1959) or an associative connection (e.g., Asratyan, 1974; Colwill & Rescorla, 1986; King, 1979; N. E. Miller, 1963; Mowrer, 1960; Pavlov, 1932; Sheffield, 1965; Sutton & Barto, 1981; see Dickinson, 1989, for a discussion of this distinction). Unless such knowledge plays a central psychological role in performance, we should have qualms about claiming that the action in question is "directed" at the goal.

Our conception of goal directedness also requires that the outcome of the action should be represented as a goal for the agent at the time of performance. If we observed that the agent readily performs the action at a time when we have independent evidence that the outcome is not in fact represented as a goal, we should again doubt whether the behavior in question is directed at this "goal."

In summary, then, an action is goal directed if its performance is mediated by the interaction of two representations: (1) a representation of the instrumental contingency between the action and the outcome, and (2) a representation of the outcome as a goal for the agent.

The status of any particular activity can be evaluated in terms of this conception of goal directedness by two empirical criteria: the *instrumental* criterion and the *goal* criterion.

Instrumental Criterion

To the extent that an action is mediated by a representation of its relation to an outcome, its acquisition and performance should be sensitive to the instrumental contingency between the action and the outcome. The aim of this criterion is to distinguish such actions from behavior mediated by a Pavlovian association between a signal or conditioned stimulus and an outcome.

Although both instrumental (Thorndike, 1911) and Pavlovian conditioning (Pavlov, 1927) were studied early in the century, it took students of learning some time to appreciate the critical difference between the two forms. S. Miller and Konorski (1969) are usually credited with being the first to make the distinction in 1928 by demonstrating the conditioning of an action, flexion of a dog's rear leg, to a signal for food that appeared to be unrelated to the response elicited by the food. This, they argued, was at variance with Pavlov's principle of stimulus substitution, according to which exposure to stimulus-outcome pairings endows the stimulus with the capacity to act as a substitute or surrogate for the outcome and thereby enables it to elicit the same responses. As the ability of the signal to control leg flexion could not be explained in terms of its becoming a substitute for food, Miller and Konorski argued for a second, Type II, conditioning process. Four years later, Skinner (1932) also proposed a Type II form of conditioning for much the same reason, namely, that Pavlov's principle could not explain why hungry rats learned to press a freely available lever for food.

Although these studies clearly provide a challenge to stimulus substitution as a universal principle of conditioning, what they did not demonstrate was the instrumental character of Type II conditioning—that it is controlled by the relation between the conditioned action and the outcome. That behavior can be controlled by this relation was first demonstrated by Grindley in a paper also published in 1932. Grindley trained restrained guinea pigs to turn their heads either to the left or the right and then back again to a central position when a buzzer sounded, in order to receive the opportunity to nibble a carrot. What established that this behavior was under the control of the action-outcome relation was the fact that the animals would reverse the direction of their head turns when the instrumental contingency was reversed. Thus, when the stimulus-outcome association between the buzzer and the carrot was kept constant,¹ the behavior was controlled by its relation to the outcome. Other things being equal, such *bidirectional* conditioning can be taken as the critical assay of instrumental control.

Traditionally, behavioral psychologists have identified types of conditioning in terms of the experimental procedure rather than in terms of the relations that actually control performance. This may well be satisfactory for many purposes but, as we shall present evidence that the motivational processes engaged by stimulus- and action-outcome relations differ, it is critical that we determine the nature of the controlling relation. Skinner (1932) was probably right when he argued that free-operant lever-

pressing by rats is controlled by the instrumental contingency. Perhaps the best evidence for this claim comes from a study of punishment by Bolles, Holtz, Dunn, and Hill (1980), who trained rats to both press down and push up a lever concurrently for a food reward. The schedule was such that sometimes a press was required for the next reward and sometimes a push, in a manner that was unpredictable to the animal. The animals learned to intersperse presses and pushes. Bolles et al. then attempted to punish one category of this bidirectional behavior by following either presses or pushes with a shock. Although the introduction of the punishment contingency suppressed both actions to a certain extent, the category upon which the shock was contingent was performed at a significantly lower level. By implementing this bidirectional assay, Bolles et al. were able to demonstrate that these actions are sensitive to their instrumental consequences.

In contrast to free-operant leverpressing, the instrumental status of other widely used behavioral tasks is ambiguous. For example, locomotion in runways and mazes has been extensively used in the study of motivation and is often characterized as instrumental, although it is readily explicable in terms of the Pavlovian control of approach to stimuli associated with the outcome. Hershberger (1986) assessed the bidirectional character of performance in a runway by reversing the normal relationship between spatially directed locomotion and access to a food outcome for chicks. He employed a "looking glass" runway, in which a food bowl receded twice as fast as a chick ran toward it and drew near twice as fast as the chick ran away from it. To the extent that locomotion is controlled by its instrumental relation to spatial translation, the reversal of the normal contingency should have presented no major problems to the animals. By contrast, a purely Pavlovian animal, being insensitive to the consequences of its actions, should never be able to adapt to such a "looking glass" world. For as long as the food bowl remains a signal for food, the animal should continue to perform the response elicited by such signals, namely, attempted approach. And this was in fact the behavior pattern observed by Hershberger—the chicks failed to learn consistently to run away from the food bowl across 100 min of exposure to the reversed contingency.

In summary, application of the instrumental criterion identifies free-operant leverpressing by rats as a potential goal-directed action, while raising serious doubts about the instrumental status of other activities, such as spatially directed locomotion.

Goal Criterion

The second, goal, criterion requires that the performance of an action should depend upon whether or not the outcome is currently a goal for the agent. In the conditioning laboratory, the standard way of determining the goal dependency of an action is to conduct an outcome devaluation test. Adams and Dickinson (1981) trained hungry rats to press a lever by using two types of food pellets with different flavors. One, the positive outcome,

was delivered contingent upon leverpressing (positive contingency), whereas to receive the other, negative outcome, the animals had to refrain from pressing for a short period (negative contingency). Clearly, the obvious goal of leverpressing was access to the positive outcome. To investigate whether this was so, Adams and Dickinson devalued this outcome for one group so that it should have no longer acted as an effective goal for these animals. Subsequently, the propensity of the animals in this group to press the lever was compared with that of a second group for which the negative outcome was devalued.

The devaluation procedure capitalized on the fact that if consumption of a flavored food is followed by gastric illness, induced in the present case by an injection of lithium chloride (LiCl), animals develop an aversion to that food so that it will no longer function as an effective goal. Thus, immediately after instrumental training, an aversion to the positive outcome was established for one group and to the negative outcome for a second group. It is important to note that the lever was not present during this aversion conditioning and that the food pellets were presented independently of any instrumental action.

To assess the effect of this devaluation treatment, the animals were once again given access to the lever, and their propensity to press was recorded. If access to the positive outcome was established as the goal of leverpressing during instrumental training, and if, as a result of the aversion treatment, this event was no longer represented as a valued goal, the animals for which this outcome was devalued should have pressed less than those for which the aversion was conditioned to the negative outcome. This is just what Adams and Dickinson (1981) observed, a result confirmed by Colwill and Rescorla (1985) in a choice procedure.

Two features of this devaluation procedure are noteworthy. The first relates to the possible role of Pavlovian processes in the effect. By equating the presentations of the positive and negative outcomes during instrumental training, Adams and Dickinson (1981) ensured that the two outcomes should have entered into equivalent Pavlovian associations with the contextual cues. As a result, the differential effect of devaluing the positive and negative outcomes must have been mediated by the instrumental rather than Pavlovian relations present during training. Controlling for the role of Pavlovian conditioning in the devaluation effect is not just a pedantic feature of the design, for a number of theorists (e.g., Bindra, 1972, 1978; Konorski, 1967; Rescorla & Solomon, 1967; Spence, 1956; Trapold & Overmier, 1972) have argued that effects of the incentive properties of an outcome, such as its magnitude and quality, on instrumental performance are mediated by a Pavlovian process. To the extent that a devaluation effect is mediated by such a Pavlovian process, we should not use this effect as evidence that the outcome functions as a goal of an action within our strict definition of goal directedness.

Second, it is important to note that the test was conducted in "extinction"; neither outcome was actually presented during the test. If Adams and Dickinson had actually presented the food pellets contingent upon leverpressing during the test, this devaluation effect could have been explained in terms of the direct suppressive effects of the now-aversive positive outcome. By testing in extinction, however, they ensured that the differential performance of the two groups must have reflected the interaction of knowledge of the instrumental contingencies acquired during initial training with some representation of the status of the outcome as a goal following the devaluation treatment.

In conclusion, the application of our two criteria, the instrumental and goal criteria, have allowed us to identify at least one goal-directed action—free-operant leverpressing by rats. As far as we know, no other activity has been demonstrated to be goal directed by these criteria and, for this reason, our subsequent discussion of the motivational control of goal-directed behavior will be primarily based upon empirical work that has employed this action. Application of the criteria for goal directedness means that we shall neglect a large body of empirical work on motivational control, such as that employing spatially directed behavior in mazes and runways, because its status as a goal-directed action is at present ambiguous. Specifically, what remains uncertain is the extent to which such behavior meets the instrumental criterion; spatially oriented behavior may well be controlled by the Pavlovian relationships between distal stimuli and the reward rather than by the instrumental action-outcome contingency. Our selection is justified not only by our prime concern with goal-directed action, but also because, as we shall see, there are reasons for believing that the motivational influences mediated by Pavlovian associations may well differ from those engaged by instrumental contingencies. Reference to the nonoperant literature will be restricted to cases in which we have found precedents for our theoretical and empirical claims.

TOLMANIAN CATHEXES

Given our psychological definition of goal-directed behavior, the problem of motivational control is one of specifying the processes by which motivational states determine the agent's representation of the outcome as a goal. In the case of the flavor aversion employed in the outcome devaluation procedure, Garcia (1989) has argued that this form of conditioning sets up a conditioned feedback loop that is activated when the flavor is subsequently contacted to produce disgust or distaste, a state that Rozin and Fallon (1987), among others, have identified as a primary emotion. At issue in this specific case, then, is the mechanism by which this state of disgust determines the goal status of the food outcome and, within this context, it is significant that Garcia explicitly relates his ac-

count of flavor-aversion learning to Tolman's concept of "cathexes." According to Tolman, a cathexis refers to a connection between a potential goal and the relevant motivational state, a connection that, in many cases, Tolman assumes is learned:

By the learning of a cathexis I shall mean, then, the acquisition of a connection between a given goal-object or disturbance-object—i.e., a given type of food, a given type of drink, a given type of sex-object or a given type of fear object—and the corresponding drive of hunger, thirst, sex or fright. (Tolman, 1949b, p. 144)

According to this account, then, the cathexis mediating the outcome devaluation effect is a connection between the experience of the flavor and a state of disgust.

Tolman was also explicit about the conditions required for the acquisition of a cathexis; cathexes are acquired through consummatory contact with the goal object:

It would seem that animals or human beings acquire positive cathexes for new foods, drinks, sex-objects, etc., by trying out the corresponding consummatory responses upon such objects and finding that they work—that, in short, the consummatory reactions to these new objects do reduce the corresponding drives (Tolman, 1949b, p. 146),

and presumably that "they don't work" in the case of negative cathexes. Although Tolman was not entirely clear about the processes underlying the acquisition of negative cathexes (see Tolman, 1949b, pp. 147–148), a generalization of his consummatory principle to the negative case suggests that devaluation by flavor-aversion conditioning involves two distinct processes. The first is a *latent* change in the affective or drive state elicited by the flavor (and possible other stimulus properties of the food or fluid), which is brought about by its pairing with gastric malaise. This change is not manifest behaviorally, however, until the agent subsequently recontacts the flavor, at which point disgust will be experienced. It is this experience of the flavor in conjunction with the state of disgust that instantiates the conditions for the second process, namely, the formation of a connection between the flavor and the negative affect of disgust, which, in Tolman's terms, represents the acquisition of a negative cathexis. Just as Tolman indicates that positive cathexes are reinforced by pairing exposure to the goal object with a reduction in the relevant motivational state, so we could argue that the pairing of such exposure with the induction of an aversive motivational state reinforces negative cathexes.

Tolman also never clearly specified the process by which a negative cathexis controls instrumental performance. In a companion paper (Tolman, 1949a), however, he outlined the mechanism by which positive cathexes are assumed to act:

By a cathexis I shall not mean the resultant loading of a given type of goal-object by drive (this loading I shall call value, see below) but rather the innate or acquired connection (or channel) between the given drive, say a viscerogenic hunger (such as food-hunger or thirst-hunger, or sex-

hunger) and the given *type* of goal-object. The result of such a cathexis (i.e., such a channel) will be that when the given channel is deep, that is highly permeable, there will be a flow of energy from the drive to the compartment for the given type of goal-object. And this will result in what I call a corresponding *positive value* for that *type* of goal-object. (Tolman, 1949a, pp. 360–361; italics are Tolman's)

Although a clear interpretation of the distinction between value and cathexis is somewhat obscured by Tolman's florid description, the gist of the distinction is probably best illustrated by casting his account into the associative metaphor. Figure 1 illustrates the associative structures assumed to be formed during an outcome devaluation procedure such as that employed by Adams and Dickinson (1981). During the first, instrumental, training stage, two associations are formed (top panel). The first is that produced by instrumental learning and takes the form of an association from a unit excited by the presentation of the food outcome to one whose activation produces leverpressing. This association corresponds to Tolman's "means-end readiness" that, when excited, represents an "expectancy."²

The second and, in the present context, most important association is the positive cathexis (the "deep" and "permeable" "channel") from the hunger system to the food unit ("the compartment for a given type of goal-object") brought about by eating the particular food used as the outcome while the hunger system is activated by food deprivation. Given the formation of these two connections, chronic activation of the hunger system by food deprivation will lead to excitation of the food unit via the cathexis, which in turn will activate the response unit to produce leverpressing when the stimulus support for this action is available. Within this characterization of the Tolmanian system, it is the activation level of the food unit that represents the current *value* of this outcome and hence its capacity to act as a goal of instrumental action, because it is via this variable that the degree of hunger is assumed to affect behavioral output.

The *first* pairing of food consumption with gastric illness during aversion conditioning adds one further connection to this structure, a connection between the *input* to the food unit and a disgust system (see middle panel of Figure 1). As the disgust system is assumed to have an inhibitory influence on units activated by gustatory stimuli, this connection opens the negative-feedback loop on the food unit identified by Garcia (1989). Consequently, when the food is re-presented (as in the bottom panel of Figure 1), activation of the food unit by the direct input will be inhibited by the disgust system through the conditioned negative-feedback loop, thus producing a suppression of intake and a loss of goal value. It is important to note, however, that this loop will not be engaged if the animal is tested for leverpressing in extinction immediately after a *single* pairing of food consumption and gastric malaise. In the absence of any activation of the *external* input to the food unit (because the food is not presented during the extinction test), the negative-feedback

Tolmanian Cathexes

Incentive value = activity of **FOOD**

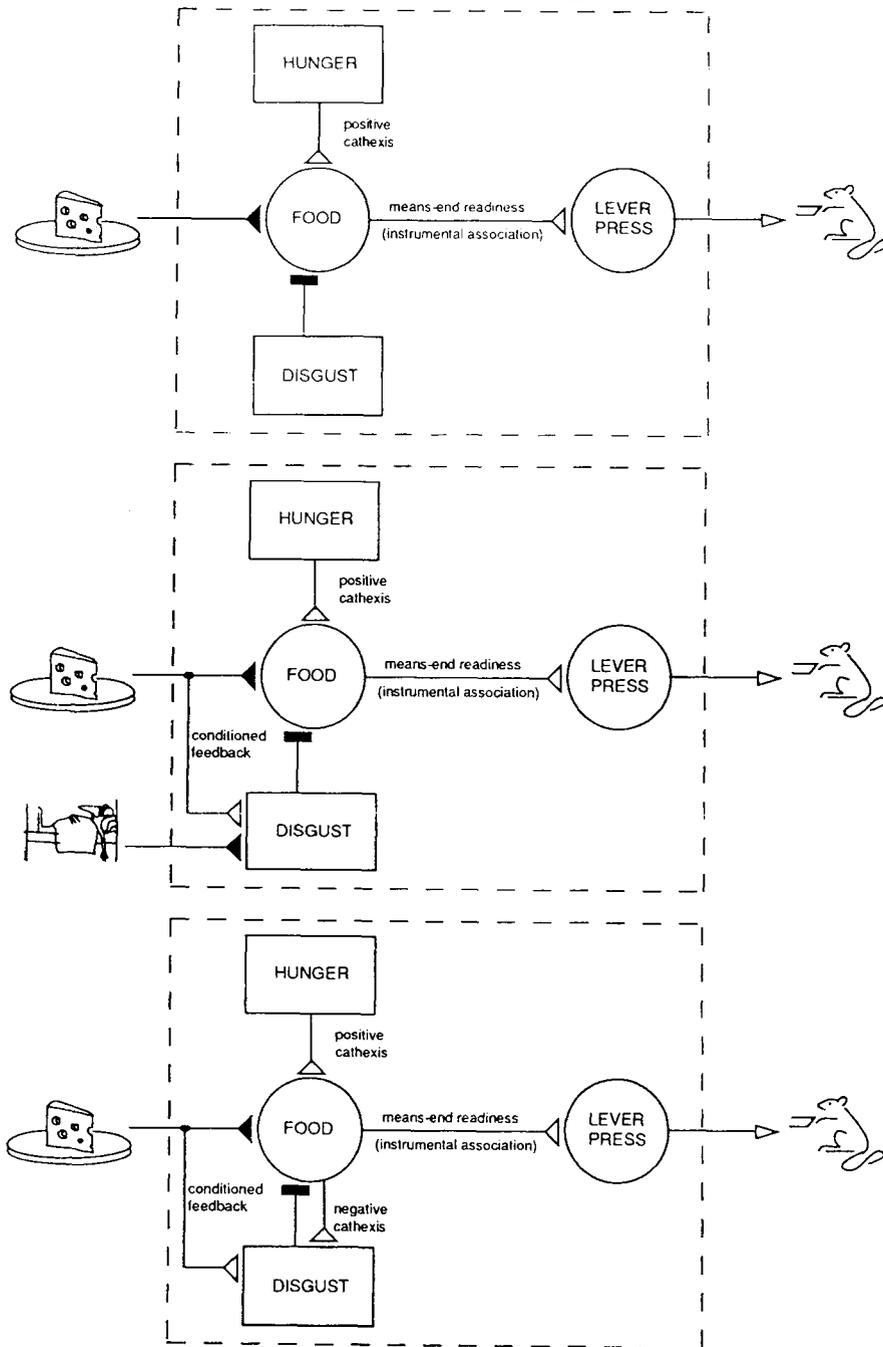


Figure 1. A cartoon illustrating the application of cathexis theory to an outcome devaluation procedure. The top panel shows the instrumental association and positive cathexis established during instrumental training with a food reward (the plate of cheese), the middle panel shows the conditioned feedback established by pairing the food with gastric malaise (a sick rat in bed), and the bottom panel the negative cathexis resulting from contact with the food after aversion conditioning.

loop will not be engaged and the excited hunger system will still elicit leverpressing as normal via its *internal* input (the positive cathexis) to the food unit.

According to cathexis theory, what is required for an outcome devaluation effect is a further exposure to the food. As we have already noted, re-presentation of the food will not only excite its unit, but also will activate the conditioned feedback loop and hence the disgust system, with the resultant suppression of intake. More important for the present purposes, however, is the fact that, according to Tolman, reexposure to the food ("trying out the corresponding consummatory responses upon such objects") should lead to the formation of a cathexis or connection between the food unit and the disgust system, presumably due to their contiguous activation. In the bottom panel of Figure 1, we have characterized this connection as an excitatory link from the food unit to the disgust system, and it is this connection that is crucial to the outcome devaluation effect. Now when the animal is tested for leverpressing, the activation level of the food unit under the excitatory influence of the hunger system will be limited by the inhibitory feedback influence from the disgust system that has been brought into play by the negative cathexis, thus reducing the tendency to perform the instrumental response.³

The aversion conditioning procedure used in the Adams and Dickinson (1981) outcome devaluation study allowed for the formation of the negative cathexis. Although robust aversions can be acquired with a single conditioning episode when the flavor is novel, the familiarization of the flavor during instrumental training probably retarded the rate of aversion conditioning so that a number of pairings of the food outcome and the induction of gastric malaise were required to suppress intake. Thus, the fact that the animals were exposed to the food on later pairings, when they should have had at least a partial aversion, provided the opportunity for the formation of the negative cathexis. As we have already noted, however, if an effective aversion could be conditioned with a single pairing of the outcome and toxicosis, this conditioning should not produce an outcome devaluation effect. In the absence of contact with the food after aversion conditioning, the negative cathexis would not be formed and performance should be controlled solely by the positive cathexis between the hunger system and the food unit acquired during instrumental training (see Figure 1).

In testing this prediction, we (Balleine & Dickinson, 1991) overcame the problem of familiarization by training our rats to press the lever for a sucrose solution in a single session. An aversion was then conditioned to this solution by injecting the experimental animals with LiCl immediately following this session, when the sucrose was novel. Although this treatment was sufficient to abolish the rewarding properties of the sucrose when it was later presented contingent on leverpressing in a reacquisition test, it had no impact upon performance in a prior extinction test. These animals pressed the lever just as frequently in this test as did control rats for which the LiCl

injection was sufficiently delayed to prevent conditioning of an aversion.

The absence of an outcome devaluation effect accords with the prediction of cathexis theory; because the experimental rats were not reexposed to the sucrose solution following aversion conditioning, there was no opportunity for the formation of the negative cathexis between the sucrose unit and the disgust system. Of course, given an opportunity for cathexis learning, the aversion treatment should affect subsequent instrumental performance. This is just what we found; if the experimental rats were simply reexposed to the sucrose between aversion conditioning and the extinction test, then their instrumental performance in this test was reduced. It should be noted that this reexposure was effective whether it occurred in the training chamber with the levers withdrawn or in a different drinking cage.

In a second study, we (Balleine & Dickinson, 1991) investigated whether reexposure to an outcome after aversion conditioning also affects instrumental choice. These rats were trained to perform two different actions, leverpressing and chain pulling, for different outcomes in a single session, which was immediately followed by LiCl injection. Presumably, this procedure should have conditioned an aversion to both these outcomes. The animals were then given the opportunity to acquire a negative cathexis for only one of the outcomes by being exposed to noncontingent presentations of this outcome in the absence of the lever and chain. In accord with cathexis theory, the action trained with the reexposed outcome was performed less frequently than the alternative in a subsequent extinction test.

Whatever the merits of this particular account of the outcome devaluation effect, the model does make clear one surprising prediction of Tolman's cathexis theory concerning the effects of simple motivational shifts. This prediction can be illustrated by a case in which a sated rat is exposed to an instrumental contingency between leverpressing and a food outcome. Such training should result in the formation of the instrumental association illustrated in Figure 1, but not the positive cathexis. Because the animal is not hungry at the time it receives the food outcome, the conditions for the formation of the connection from the hunger system to the food unit will not hold. And, in the absence of this connection, a shift to the state of hunger (activation of the hunger system) should have no impact on the value of the food (excitation level of the food unit) and hence on instrumental performance. This prediction only holds, of course, when the instrumental performance is tested in extinction to preclude exposure to the food under hunger and hence the formation of the cathexis.

In spite of the extensive literature on the effect of motivational shifts on instrumental performance (see Bolles, 1975), we have found only two free-operant studies that bear on this prediction of cathexis theory. In the first, Sackett (1939) trained rats, maintained on a 23-h food-deprivation schedule, to press a lever for food pel-

lets in a single session given 7 h after the daily feeding period. When subsequently tested in extinction, the rats tended to press more if tested 30 rather than 6 h after the feeding period. Although this difference was not reliable in Sackett's study, essentially the same pattern was subsequently reported by Crocetti (1962).

There are, however, two reasons for arguing that such a result does not provide an adequate assessment of cathexis theory. The first, and most obvious, is that the level of deprivation during training may well have been sufficient to have established a connection between the hunger system and the food unit so that increasing activation in the hunger system would have enhanced the excitation of the food element and hence the value of this incentive. Moreover, a close inspection of the procedures in these studies reveals a second problem, in that the training food reward in both cases consisted of a dried pellet molded from the lab chow used for the daily maintenance of the animals. Consequently, the animals had experienced this food while hungry during their daily maintenance and thus had had the opportunity to learn the appropriate cathexis between this food and the motivational state induced by food deprivation.

In the absence of any adequate test of cathexis theory, Balleine (1992) reexamined the effects of shifts in motivational state using unfamiliar foods, either the standard high-protein Noyes reward pellets or a polysaccharide (maltodextrin) solution, as the outcome. In one study, the rats were trained to press a lever for the food outcome in the absence of any deprivation of maintenance diet in the home cage before performance was tested in extinction either in the same, undeprived state or following 22.5 h of food deprivation. Surprisingly, the deprivational state at the time of testing had no effect on performance, which was uniformly low in both conditions. This is just the result anticipated by Tolman's cathexis theory. In the absence of any prior experience with the Noyes pellets or maltodextrin in the deprived state, the animals had no opportunity to acquire the appropriate cathexis between these outcomes and the state of hunger. Consequently, inducing this state prior to the extinction test should not have enhanced the incentive value assigned to these outcomes.

Of course, instrumental performance should be sensitive to the induction of hunger following experience with the relevant food outcome in this state, an experience that would allow the animals to acquire the appropriate cathexis. A T-maze study by Wetzel (1959) provides partial evidence for such cathexis learning by comparing the acquisition of a discrimination between a sucrose solution and water by hungry and nondeprived rats. The performance of the hungry rats was superior to that of the nondeprived animals only for a group that had separate experience with consuming the sucrose solution while hungry. Unfortunately, Wetzel did not include a condition in which animals had separate exposure to the sucrose in the nondeprived state, so we cannot be certain that this effect was due to exposure to the sucrose when hungry.

To examine the effect of giving the animals the opportunity to acquire the appropriate cathexis for the food outcomes under hunger, Balleine (1992) ran a third group that was allowed to consume the relevant food in separate cages while food deprived prior to any instrumental training. A very different pattern of results emerged; following instrumental training in the nondeprived state, the induction of hunger elevated the rate of leverpressing in the extinction test relative to both the nonexposed groups and a fourth group that received preexposure to the food outcome when hungry but was tested in the nondeprived state. In a second experiment, Balleine demonstrated that the opportunity for cathexis learning also affects the choice between alternative instrumental actions. Nondeprived rats were taught that one action, leverpressing or chain pulling, produced one outcome, Noyes pellets or maltodextrin solution, and that the other action produced the alternative outcome. Prior to this training, however, all the animals had received access to one of the outcomes when nondeprived and the other following food deprivation. When the animals were given a choice between the two actions in extinction under hunger, they showed a reliable preference for the action associated during training with the outcome that had been preexposed under hunger.

It would appear, therefore, that primary motivational states, such as hunger, do not determine the value of an instrumental goal directly; rather, animals have to learn about the value of a commodity in a particular motivational state through direct experience with it in that state, learning that is characterized within Tolman's cathexis theory by the formation of a connection from the motivational system to a representational unit for the outcome.

RESISTANCE TO SATIATION

Although cathexis theory predicts the role of outcome preexposure in the sensitivity of instrumental performance to the induction of hunger, the effects of other motivational shifts are less readily explained. We shall illustrate this problem by considering the reverse shift, namely, from training under a state of hunger to testing in a nondeprived state. The point at issue is whether cathexis theory predicts that instrumental performance should be immediately sensitive to this shift. The prediction is clear: reducing the level of hunger should produce an immediate reduction of instrumental performance. This is because, according to Tolman, outcome value is determined by the current activation level of the food element, which in turn receives its input (via the cathexis) from a system sensitive to the agent's state of hunger (see Figure 1). Reducing the level of food deprivation should consequently produce an immediate loss of value for the food, which in turn should be reflected in a decrement in instrumental performance. And indeed there are a number of operant studies that have demonstrated that a reduction in primary motivational state between training and testing produces a corresponding reduction of instrumental performance during an extinction test (Crocetti, 1962; Eis-

man, Theios, & Linton, 1961; Grice & Davis, 1957; Heathers & Arakelian, 1941; Horenstein, 1951; Koch & Daniel, 1945; Perin, 1942; Saltzman & Koch, 1948; Skinner, 1936).

There is, however, a body of evidence on the so-called "resistance-to-satiation" effect suggesting that agents may have to learn that a previously trained instrumental action no longer yields a valuable outcome when tested in a sated or semisated state. As Morgan (1974) notes in a review of this topic, a number of studies have shown a progressive rather than immediate decline in rewarded performance when the agent is tested in a sated or semisated state after training in a deprived state, a finding that he interprets as evidence that animals have to learn that the outcome is no longer valuable in the sated state. Direct evidence for such learning comes from a runway study by Peterson and McHose (1980). In this study, a decrement in running speed was observed following a shift from a high to low level of food deprivation only if the rats had previously had the opportunity to consume the food reward while in the low-deprivation state.

As in the case of the upshift studies, a possible resolution between the resistance-to-satiation effect and the immediate effect of a downshift in motivational level on extinction performance may lie with the animals' previous experience with the instrumental reward or outcome. Except in one case, all studies that reported an immediate effect of a downshift on extinction performance employed an instrumental reward made up from the animals' maintenance or satiation diet. Consequently, in these studies, the animals had an opportunity to learn about the reduced value of the outcome in the low motivational state prior to testing. In the remaining report (Eisman et al., 1961), the composition of the reward is not described.

In summary, it is unclear whether or not a reduction in primary motivational level produces an immediate attenuation of operant performance in the absence of any prior experience with the incentive in the reduced state. Consequently, Balleine (1992) examined this issue by training hungry rats to press a lever for a novel incentive, either the Noyes pellets or the maltodextrin solution, before testing performance in extinction when the animals were either hungry or not deprived of food. The deprivational state on test had no detectable effect on instrumental performance; the nondeprived animals pressed just as vigorously as the hungry ones, a finding clearly at variance with the prediction of cathexis theory.

To investigate whether prior experience with an outcome in the nondeprived state affects sensitivity to a reduction in motivation, Balleine (1992) gave two further groups exposure to the outcome in this state prior to training under hunger. A very different pattern of results emerged when these animals were tested in extinction; the nondeprived animals pressed much less often than the hungry ones, for which performance in turn was unaffected by the preexposure treatment. As in the case of motivational upshifts, Balleine found that this treatment also affected choice. In this study, the animals were pre-

exposed to one outcome when hungry and the other when nondeprived prior to concurrent training with both outcomes under hunger. When these animals were tested under extinction in the nondeprived state, they showed a preference for the action that, during training, had yielded the outcome that had been preexposed under hunger. In other words, preexposure in the nondeprived state appeared to reduce the relative value of an outcome when that state was subsequently induced.

Just as agents appear to have to learn about the positive outcome value of an outcome in a relevant motivational state, so they also have to learn about the low value of the outcome in the absence of that state. Although cathexis theory anticipates the former results, the resistance to satiation in the absence of prior exposure in the nondeprived state is clearly problematic for Tolman's account of the control of outcome value by motivational variables. There is now increasing evidence for the generality of persistence of instrumental performance in the face of reductions in motivational level and for the role of incentive exposure in mediating adaptation to such shifts. Lopez, Balleine, and Dickinson (1992a) found a comparable dependency in the case of a transition from training under thirst to testing in a nondeprived state both for the absolute frequency of pressing and for choice. Everitt and Stacey (1987) employed a free-operant leverpressing schedule to investigate the effect of castration on the sexual motivation of male rats. Having trained the animals to press throughout a 15-min session that terminated with access to an estrous female, the male rats were castrated and allowed a 40-day recovery period before testing. Although the castration produced an immediate reduction in a number of unconditioned sexual behaviors elicited by the first postoperative presentation of the female at the end of the initial test session, there was no detectable effect of the operation on instrumental performance prior to this presentation. Presumably, in absence of exposure to a female in the castrated state, instrumental performance was controlled by the high value assigned to this outcome during training. The exposure to the female at the end of the first test session, however, should have allowed the castrated rats to learn about the lower value of the female in this state, and indeed these animals pressed at a significantly lower rate than controls on all subsequent test sessions.

INCENTIVE LEARNING

These data lead us to the conclusion that primary motivational states have no direct impact on the current value that an agent assigns to a past outcome of an instrumental action; rather, it appears that agents have to learn about the value of an outcome through direct experience with it, a process that we refer to as *incentive learning*. As a result, certain operations that potentially affect the value of an outcome, such as motivational shifts and aversion conditioning, only do so once the agent has the opportunity for incentive learning by reexperiencing the outcome in the altered state.

This form of incentive learning must be clearly distinguished from the learning processes by which outcome value, once acquired, comes to control instrumental performance. This distinction can be illustrated by reconsidering the cathexis model illustrated in Figure 1; whereas incentive learning within the context of this model refers to the formation of the connection between the hunger system and the food unit, outcome value (the activation level of the food or outcome unit) controls performance via the instrumental association between outcome and action units. Moreover, as we have already noted, others have suggested that incentive properties of an outcome can affect instrumental performance through Pavlovian conditioning to contextual and discriminative stimuli (indeed, we argue below for such a process) and, as a result, have characterized such stimuli as incentive stimuli. This Pavlovian process is to be distinguished, however, from incentive learning, which refers to learning about the value of the outcome itself.

Although our concept of incentive learning is within the general spirit of Tolman's cathexis theory, aspects of the process appear to be at variance with the details of his account. Specifically, the role of incentive learning in mediating the adjustment of instrumental performance to reductions of motivational level are not anticipated by cathexis theory. The problem with Tolman's theory lies with its failure to differentiate between two functions of motivational states in the control of instrumental action. The first is the determination of the assigned value; contact with a food under hunger results in the assignment of a higher value than exposure when sated. This assignment is not automatically made conditional upon the motivational state during exposure, however. This is evident from the fact that, in the absence of prior outcome exposure under the test state, rats persist at performing at a level determined by their training state. Thus, for example, when animals were trained either hungry (Balleine, 1992) or thirsty (Lopez et al., 1992a), performance during the extinction test was unaffected by whether the animals were deprived or sated during the test. In other words, the high value assigned during training in the deprived state continued to determine performance in the sated state.

The idea that the motivational state at the time of contact endows a commodity with an incentive value that generalizes across states has a precedent in studies by Revusky (1967, 1968). In these experiments, rats were given the opportunity to consume one flavor in a high deprivational state and another in a low state before preference for the two flavors was assessed. Revusky found that the flavor consumed in the high state was preferred, regardless of the deprivational conditions at the time of testing. Although it is disputed whether Revusky's effect was actually due to the deprivational state at the time of consumption (Capaldi & Myers, 1982; Capaldi, Myers, Campbell, & Sheffer, 1983), Capaldi, Sheffer, and Owens (1991) have subsequently confirmed the generalized preference for unsweetened food eaten while hungry. In-

deed, Capaldi and her colleagues (Capaldi, Davidson, & Myers, 1981) have demonstrated that prior exposure to a food under hunger establishes it as an effective reinforcer for runway performance when the animals are sated.

The second function of a motivational state is the acquisition of control over the assigned value so that performance is sensitive to shifts in state. Thus, the animals that were sensitive to a downshift in motivational state in the Balleine (1992) and Lopez et al. (1992a, 1992b) studies experienced the outcomes both when sated in the initial exposure stage and when hungry during instrumental training. Presumably, this experience allowed the animals to learn that the outcome had a low value when sated and a higher value when deprived. In other words, this exposure had enabled the motivational state of the animal to acquire control over the assigned value of the outcome.

In summary, the motivational state of an agent appears to play two distinct but interacting roles in incentive learning. The first is the *determination* of the assigned value at the time of contact with the outcome, and the second reflects the acquisition of *control* over the assigned value by the current motivational state. We shall now consider the possible processes mediating each of these roles.

Determination of Outcome Value

At present, we know next to nothing about the psychological processes underlying incentive learning. An obvious candidate is learning about the feedback consequences of contact with or consumption of the outcome. Thus, in the case of devaluation by flavor-aversion conditioning, the agent may learn that contact with or consumption of the outcome predicts nausea and disgust. Similarly, learning about the nutritional and hydrational consequences of an outcome when hungry, thirsty, or replete may mediate the determination of value by motivational states. Indeed, such learning would appear to be what Tolman (1949b) had in mind when he claimed that the acquisition of cathexes depends upon the agent's finding out that consummatory responses "work—that, in short, the consummatory reactions to these new objects do reduce the corresponding drives" (Tolman, 1949b, p. 146).

An alternative view of incentive learning emphasizes the hedonic properties of the outcome, with high values being assigned to outcomes experienced as hedonically attractive and low or negative values to ones experienced as noxious or unpleasant (Cabanac, 1971; Young, 1961). Within this account, incentive learning consists of assigning a value on the basis of the agent's affective or hedonic response to the outcome itself rather than on the basis of knowledge about the consequence of consummatory contact with the outcome. Interestingly, Baeyens, Eelen, van den Berg, and Crombez (1990) have recently presented evidence for just such a hedonic shift by using a flavor-aversion procedure that appears to be independent of any knowledge of the associates of the flavor. They gave human subjects both sweet and distasteful solutions to drink, each flavored with a different but distinctive taste cue. Although experience of the association between a

flavor and the distasteful solution reduced the hedonic rating of the flavor cue when subsequently presented by itself, the subjects were unable to report reliably which flavors had been paired with the sweet and distasteful solutions. This was not simply due to the insensitivity of the measure of the subject's knowledge of the flavor-solution associations; when color rather than flavor cues were used, the subjects could reliably report the color-solution pairings even though these pairings did not alter the hedonic responses to the colors themselves. Campbell, Capaldi, Sheffer, and Bradford (1988) have claimed that conditioned flavor preferences and cuing functions can also be dissociated in rats.

Of necessity, there are no direct measures of the hedonic reactions of animals, although Grill and Berridge (1985) have argued that the hedonic properties of a food or fluids can be assessed indirectly by the consummatory response patterns elicited by oral contact in rats. Attractive fluids, such as a sugar solution, elicit a specific pattern of ingestive responses, whereas aversive or defensive responses are elicited by quinine and other noxious substances. If it is accepted that these response patterns reflect the animals' affective response to a substance, then there is evidence that procedures that change the value of foods and fluids as instrumental outcomes also directly modulate the hedonic responses to these substances. Berridge, Grill, and Norgren (1981) reported that the flavor-aversion procedure used in outcome devaluation studies changes the animals' consummatory responses to the outcome. Pairing an attractive flavor, such as sucrose, with gastric illness induced by LiCl shifts the initially ingestive pattern to the aversive pattern characteristic of a distasteful substance. Moreover, Berridge and his colleagues have demonstrated that the induction of a motivational state can directly and immediately enhance an animal's hedonic response to a relevant substance. Whereas an oral infusion of hypertonic saline normally elicits a mixed pattern of ingestive and aversive responses, the first infusion when rats were sodium depleted immediately elicited exclusively the ingestive pattern (Berridge, Flynn, Schulkin, & Grill, 1984). Similarly, food deprivation enhances the ingestive patterns elicited by a quinine-adulterated sucrose solution (Berridge, 1991). These data clearly accord with the idea that motivational states act by modulating the hedonic response to a relevant outcome that, when experienced, leads to the assignment of an outcome or goal value based upon this response.

This is not to claim, of course, that knowledge of the consequences of exposure to or contact with an outcome does not change its value. In fact, the study of consummatory response patterns engendered by different aversion procedures indicates the circumstance in which such knowledge may well play a role. A number of years ago, Pelchat, Grill, Rozin, and Jacobs (1983) reported that although the consumption of sucrose could easily be suppressed by pairing it with electric shock, this suppression was not accompanied by a shift in the positive ingestive response elicited by the solution. The rats still showed

the full pattern of ingestive responses on contact with the sucrose solution, even though they drank little of it. This was in contrast to the case in which consumption was suppressed by LiCl, which, in agreement with Berridge et al. (1981), produced a shift to the aversive response pattern. Pelchat et al. suggested that, whereas the suppression of intake conditioned by gastric malaise reflects a shift in the hedonic response to the sucrose itself, that conditioned by somatic discomfort is due to avoidance of consumption based upon knowledge of its association with the aversive consequence.

The idea that outcome value can be determined by both hedonic responses and predictive learning provides a possible resolution of an apparent inconsistency in the outcome devaluation literature. The hedonic account, like Tolman's cathexis theory, anticipates that an outcome devaluation effect should not be observed after a single pairing of the outcome and an aversive reinforcer, a claim in accord with Balleine and Dickinson's (1991) failure to detect such an effect without reexposure to the outcome following aversion conditioning. In the absence of reexposure, the animal has no opportunity to learn about its changed reactions to the outcome, which is the basis for a reassignment of value according to the hedonic theory. At variance with this prediction, however, is a recent report by Rescorla (1992), who used a choice procedure in which rats were trained to perform two different actions for different outcomes. One of the outcomes was then devalued by a *single* pairing of its consumption with an injection of LiCl before the animals' preference for performing the two actions was tested in extinction. Even though the animals were not reexposed to the devalued outcome prior to the choice test, they were less inclined to perform the action that had produced this outcome during training. In the absence of any opportunity for incentive learning, it would appear that Rescorla's devaluation effect must have been mediated by learning about the predictive relation between the outcome and the consequences of the LiCl injection at the time of conditioning.

We surmised that the critical difference between the Balleine and Dickinson (1991) and Rescorla (1992) procedures was the tonicity of the LiCl solution used to condition the aversion. Although the dose was the same in the two studies, Rescorla administered it as a hypertonic solution, whereas we employed an isotonic injection. Informal observation suggested that a hypertonic solution produces peritoneal discomfort at the time of injection, raising the possibility that the LiCl injection employed by Rescorla may have acted, at least in part, like the shock in the Pelchat et al. (1983) study and brought about devaluation through learning about the association between the outcome and somatic discomfort.

To investigate this possibility, we (Balleine & Dickinson, 1992) repeated our devaluation procedure by employing a hypertonic LiCl injection, and replicated Rescorla's devaluation effect without reexposure. If, however, the discomfort produced by the hypertonic injection was ameliorated by administering it under a short-term anesthetic,

reexposure was once again found to be necessary for the emergence of a devaluation effect. Thus, it would appear that aversion conditioning can devalue an outcome by two different processes. First, associating the outcome with somatic discomfort brings about devaluation by establishing the outcome as a signal for a noxious event. By contrast, pairing the outcome with gastric malaise alone requires that the agent has the opportunity for incentive learning if this procedure is to change the outcome value.

Control of Outcome Value

Whatever the merits of these accounts of how motivational states determine outcome value, they still leave open the question of how such states acquire control over value so that goal-directed behavior adjusts appropriately to shifts in state. The most straightforward account would argue that motivational states act like any other predictive stimuli, such as Pavlovian and discriminative stimuli, by being presented in association with the predicted state of affairs. But in the case of incentive learning, the association would be between the motivational state and the value of the outcome, which is in turn determined by that state. Traditional drive theory has long recognized the stimulus as well as motivational function of motivational states (Webb, 1955). Moreover, just as the acquisition of discriminative control often depends upon differential reinforcement (e.g., Jenkins & Harrison, 1960), so motivational control over outcome value appears to depend upon experience of differential values in the presence and absence of the motivational state.

Normally, of course, the outcome—for example, a food reward—would have a high value under the relevant motivational state—in this case, hunger—and a relatively lower value in the absence of the state, so that the value-determining and discriminative functions of the motivational state are congruent. Capaldi, Viveiros, and Davidson (1981) illustrated the interaction of the two functions of hunger, however, by reversing the normal relationship between motivational state and outcome value. They trained rats to use either a high or low level of food deprivation to predict the occurrence of a food reward in a runway. For one group (H+/L-), the two functions were congruent in that a high level of hunger predicted food and a low level predicted nonreward. For the second group (L+/H), by contrast, it was the low level of hunger that predicted the food reward so the higher outcome value was predicted by the lower motivational state. Although the performance of both groups came under the control of the motivational state, with the animals running faster in the rewarded state, a comparison of the running speed of the two groups in the low state yielded an apparently paradoxical result. The animals for which this state predicted reward (Group L+/H-) ran more slowly, at least during the initial stages of training, than those (Group H+/L-) for which the low state predicted nonreward. This, Capaldi et al. (1981) argued, was because Group H+/L- received the food reward under the high level of deprivation and thus assigned a greater value to this outcome

than did Group L+/H-, which received the food under the low state. Until discriminative control by the motivational state was established through differential reinforcement, this greater assigned value generalized across states to elevate performance under the low level of deprivation in Group H+/L-.

Basic functional considerations raise problems, however, for the thesis that motivational control of outcome value represents a straightforward example of discriminative or conditional control by motivational stimuli. It makes sense, of course, that the assignment of most positive values should be conditional upon motivational state, because the benefits associated with most primary positive outcomes, such as food and water, in fact depend upon the current state of the animal. By contrast, the deleterious effects of noxious events are typically independent of motivational state, and therefore functional considerations suggest that the assignment of negative values through incentive learning should not come under motivational control.

This point can be illustrated by reconsidering the outcome devaluation effect brought about by flavor-aversion conditioning. Recall that Balleine and Dickinson (1991) demonstrated that conditioning an aversion to an instrumental outcome via the induction of gastric malaise only reduced subsequent instrumental performance if the animal had received the opportunity to learn about the new, negative value of the outcome through reexposure to it after aversion conditioning. The question at issue is whether one would expect this assignment of value to come under the control of the motivational state at the time of reexposure. If such control is just an example of general discriminative processes, there is no reason to expect a failure of motivational control in this negative case. Indeed, Davidson and his colleagues (Davidson, 1987; Davidson, Flynn, & Jarrard, 1992) have recently demonstrated rapid acquisition of discriminative control over aversive conditioning by the degree of food deprivation.

On the other hand, it would make little functional sense in this case if the negative incentive learning failed to generalize across motivational states. Flavor-aversion conditioning is usually viewed as a behavioral defense against poisoning, and as the toxic effects of foods and fluids do not typically depend upon motivational state, it is of clear benefit for incentive learning to generalize across such states. In fact, Lopez, Balleine, and Dickinson (1992b) have recently provided direct evidence for generalization in this case. They replicated the outcome devaluation procedure of Balleine and Dickinson (1991) while varying the relationship between the motivational state during instrumental training and testing and that during reexposure to the outcome. Thus, for example, animals were trained and tested while thirsty, but received outcome reexposure while hungry. In their hands, the negative incentive learning observed with this procedure generalized fully across motivational state. With counterbalanced shifts between food and water deprivation, the size of the outcome devaluation effect was comparable, whether the animals received reexposure to the outcome and instrumental train-

ing and testing under either the same or different motivational states.

While the motivational independence of negative incentive learning makes functional sense in this case, there are other examples in which noxious properties of an incentive are dependent upon the agent's state. One such case is that of thermoregulation. And, indeed, it was by studying control of avoidance of a heat source by ambient temperature that Hendersen and Graham (1979) provided one of the first clear examples of incentive learning. In their study, rats were trained to avoid the onset of heat source in a warm environment before performance was tested in extinction when the ambient temperature was either high or low. Presumably, a shift to the lower temperature should alter the value of the heat source from negative to positive, a change that should be reflected by a reduction in the number of avoidances in extinction when the animals were tested in the cold environment. In fact, the ambient temperature had no effect on the level of test performance unless the animals had previously experienced the heat source when cold. Only the animals that had received prior noncontingent exposure to the heat lamp when cold showed a reduction in subsequent avoidance performance when tested in a low temperature. Whether or not this effect should be interpreted in terms of the motivational control of *negative* outcome value is, however, unclear. The reduction of avoidance by the animals receiving prior exposure to the lamp when cold could have been due either to learning that the lamp's negative value was conditional upon being hot or that its positive value was conditional upon being cold (or, of course, both). Until this ambiguity is resolved, we must remain agnostic about whether or not negative outcome values come under motivational control and, more generally, whether such control is just a special case of standard discriminative or conditional learning.

THE IRRELEVANT-INCENTIVE EFFECT

For those familiar with the so-called irrelevant-incentive effect, our claims about the role of incentive learning in the motivational control of instrumental action will appear at variance with a substantial body of evidence on the immediate effects of shifts between different motivational states. For example, in a classic study of sodium appetite, Kriekhaus and Wolf (1968) trained thirsty rats to press a lever for a sodium solution, water, or a solution of a nonsodium salt before testing performance in extinction under a sodium appetite. Interestingly, the animals trained with the saline pressed more on test than those trained with a nonsodium outcome, a difference that does not emerge if testing occurs either under thirst or in the absence of any explicit deprivation or appetite (Dickinson & Nicholas, 1983). The obvious interpretation of this finding is that the animals learned about the properties of the outcome produced by leverpressing during training, such as the sodium content, which were irrelevant to the motivational state in force at that time. The

induction of a sodium appetite then reduced the value of outcomes with properties irrelevant to this state while maintaining that of the saline, a difference that was then reflected in test performance.

The problem for incentive learning theory presented by the Kriekhaus and Wolf (1968) study is obvious; even though their rats had never been exposed to the outcomes under a sodium appetite, and therefore never had an opportunity to learn about their relative values in that state, the instrumental performance on test indicates that the shift to that state brought about an immediate and appropriate adjustment of relative value. Moreover, the adaptation of instrumental behavior to such motivational shifts appears to be general and not peculiar to sodium appetites. Woods, Vasselli, and Milan (1977) reported a comparable effect following training under thirst with a ferrous solution as the target outcome when their animals were tested in an anemic state. Dickinson and Dawson (1987b) trained food-deprived rats to press for either a sucrose solution or dry food pellets before testing performance in extinction under thirst or hunger. Whereas the type of training outcome had no significant effect on test performance following food deprivation, thirsty rats pressed more following training with the sucrose solution, the outcome relevant to the test state.⁴ Finally, Dickinson and Balleine (1990) reported an analogous irrelevant-incentive effect for a shift from training under thirst to testing under hunger when sucrose solution and saline acted as the training outcomes. In that study, the rats pressed more when tested hungry if the training outcome was the sucrose solution rather than saline.

We doubt, however, that these standard irrelevant-incentive effects are in fact mediated by motivational control of *outcome* values, that is, values that exert their effects through the agent's knowledge of the instrumental contingency between action and outcome. If shifts between motivational states produce an immediate change in relative outcome value, this change should be manifest in choice behavior as well as in the absolute level of performance of a single action. And yet research in our laboratory has consistently failed to detect a reliable irrelevant-incentive effect in a choice test. For example, Dickinson (1986) trained thirsty rats to press a lever and pull a chain concurrently, with one action producing saline and the other water. When subsequently tested in extinction under a sodium appetite, there was no reliable preference for the action trained with the sodium solution over that trained with water. It is not, however, that the animals were insensitive to the presentation of the saline during training. In a further study, Dickinson (1986) compared the effects of giving thirsty rats concurrent training with the sodium solution and water as outcomes with that found when a potassium rather than sodium solution was employed as the alternative outcome to water. When the animals were given a choice between the two actions in extinction under a sodium appetite, once again Dickinson failed to observe a reliable preference for the action trained with the sodium solution. The animals were sensitive,

however, to whether they had been exposed to the sodium at all during training. The animals that received the sodium during training, whether contingent upon leverpressing or chain pulling, performed both actions more on test than the animals trained with the potassium solution. What appears to be critical for the emergence of the irrelevant-incentive effect is simple exposure to the irrelevant incentive, in this case sodium, during training (see also Dickinson & Nicholas, 1983).

Essentially the same conclusion arose from studies of other irrelevant-incentive effects. Dickinson and Dawson (1987b), for example, gave hungry rats concurrent training with the lever and chain using a sucrose solution and food pellets as the outcomes. In contrast to the case in which animals are just trained with either the sucrose solution or the food pellets alone (see above), the animals trained with both outcomes did not show greater performance of the action trained with the sucrose solution when tested thirsty in extinction. Taken together, these results raise serious doubts about interpreting the standard irrelevant-incentive effect in terms of the modulation of outcome value by the current motivational state. If the reason why a thirsty animal performs an action that in the past has yielded a sucrose solution is because this state leads to the assignment of a relatively high positive value to this outcome, we should also expect it to select that action over one that had been trained with an outcome irrelevant to the state of thirst. Within this context, it should be noted that we have already seen that incentive learning treatments control not only the absolute level of performance of a single instrumental action, but also the relative performance or choice between alternative actions across shifts between different levels of the same motivational state (Balleine, 1992; Balleine & Dickinson, 1991; Lopez et al., 1992a).

In conclusion, what appears to be critical for the standard irrelevant-incentive effect is that, during training, an animal should receive simple exposure to the outcome that is relevant to the test state, the sodium solution in the case of a sodium appetite and the sucrose solution in the case of thirst. Whether or not this outcome is contingent upon the target instrumental action during training seems to be of little import.

PAVLOVIAN MOTIVATIONAL PROCESS

The reason why this is so probably relates to the role of Pavlovian conditioning in the control of instrumental performance. It has long been recognized that Pavlovian stimuli can modulate instrumental performance (Konorski, 1967; Rescorla & Solomon, 1967; Spence, 1956; Trapold & Overmier, 1972), and so the Pavlovian association between the contextual cues and the outcome during training should endow these cues with the capacity to support the instrumental performance during an extinction test (e.g., Estes, 1943). Moreover, it would not be unreasonable to assume that the level of support depends upon the relevance of the outcome presented during train-

ing to the motivational state on test. As a result, when hungry rats are trained to leverpress for either a sucrose solution or food pellets, the contextual cues associated with the sucrose solution should exert the greater Pavlovian influence when the animals are tested thirsty.

This Pavlovian-instrumental interaction thus provides a ready explanation of why the standard irrelevant-incentive effect depends simply upon exposure during training to the outcome relevant to the motivational state on test. The failure to detect an irrelevant-incentive effect with a concurrent choice procedure also follows directly from the Pavlovian account. In the case of the transition from hunger to thirst, concurrent training with both the sucrose solution and the food pellets ensures that the contextual cues are equally associated with both outcomes, whatever their assignment to the two instrumental actions. Consequently, during the choice test under thirst, the contextual cues will exert a comparable facilitation on both actions.

Dickinson and Dawson (1987a) assessed the role of Pavlovian conditioning by training hungry rats on a four-ply *multiple* schedule consisting of two Pavlovian components and two instrumental components (see Table 1). The levers were withdrawn during the Pavlovian components and the animals received noncontingent presentations of the sucrose solution in one stimulus (CSs) and free food pellets in the other (CSp). Interspersed with these stimuli were the two instrumental components during which the levers were inserted and different discriminative stimuli presented. During one stimulus (SDs) leverpressing was reinforced by the sucrose solution, and during the other (SDp) it was reinforced by the food pellets. On test, the animals' motivational state was shifted to thirst and leverpressing was measured in extinction during both the Pavlovian and the discriminative stimuli.

An irrelevant-incentive effect was observed in that, on test, the thirsty rats pressed more frequently during the discriminative stimulus (SDs) in which leverpressing had been reinforced with the sucrose solution than in the one associated with food pellets (SDp). That this effect was probably due to Pavlovian conditioning to the discriminative stimuli was indicated by the fact that a comparable difference was observed during the Pavlovian stimuli; even though the animals had never been trained to press the lever in these stimuli, they did so more during the stimulus (CSs) paired with the sucrose solution than during that (CSp) associated with the food pellets. Such a difference is not observed, however, if the animals are tested hungry (Dickinson & Dawson, 1987a, Experiment 2), indicating that this effect depends upon the relevance of the

Table 1
Design of Dickinson and Dawson's (1987a) Experiment

Training (Hunger)	Test (Thirst)
CSs → suc, CSp → pel	CSs:Lp > CSp:Lp
SDs:Lp → suc, SDs:Lp → pel	SDs:Lp > SDp:Lp

Note—CS = Pavlovian conditioned stimulus; SD = discriminative stimulus; Lp = leverpress; suc = sucrose solution; pel = food pellet.

training outcome to the motivational state on test. Dickinson and Balleine (1990) have reported a comparable Pavlovian instrumental transfer effect across a shift from thirst to hunger when sucrose and sodium solutions were employed as the outcomes.

Theories of Pavlovian instrumental transfer offer two explanations of this effect. That proposed, for example, by Spence (1956) and Rescorla and Solomon (1967) claims that Pavlovian stimuli exert a *general* motivating effect on instrumental behavior, whereas Trapold and Overmier (1972) argue for the associative or cuing role of Pavlovian conditioning. They suggest that through Pavlovian conditioning, discriminative stimuli (and presumably contextual cues) elicit an "expectancy" of the outcome, which in turn becomes part of the stimulus complex controlling instrumental performance. Thus, in the Dickinson and Dawson (1987a) study, expectancies of both the sucrose solution and the food pellets should have acquired control over leverpressing during training with the relevant discriminative stimulus. If it is assumed that the capacity of a stimulus to elicit an expectancy for a particular outcome is *directly* (i.e., without incentive learning) modulated by the relevance of the outcome to the agent's current motivational state, a ready explanation of the irrelevant-incentive effect is available. The expectancy of the sucrose solution elicited by its associated discriminative and Pavlovian stimuli should have been maintained when the animals were tested thirsty, whereas the expectancy of the food pellets would have been depressed, an imbalance favoring the cuing of leverpressing in the presence of the sucrose stimuli. An explanation in terms of the motivational properties of Pavlovian stimuli also has to assume modulation of expectancies by motivational state, but in this case an appeal is made to the general motivating influence of an activated expectancy rather than its cuing function. According to this account, thirsty rats pressed more for the stimuli associated with the sucrose solution not because these stimuli exerted greater discriminative control over this action, but because they enhanced the animals' general drive (Hull, 1943) or arousal level (see Dickinson & Dawson, 1987a, Experiment 2, for evidence of such a general motivational effect).

Whatever the merits of the motivational and cuing theories, there is no reason to believe that the instrumental contingency makes any detectable contribution to the standard irrelevant-incentive effect; the effect can be adequately explained in terms of the properties acquired by contextual cues and discriminative stimuli as a result of their Pavlovian associations with the outcomes. Two important theoretical conclusions arise from this analysis. The first is that simple shifts between different motivational states do not directly affect performance by modulation of the value of the outcome as a goal of the instrumental action and therefore do not represent a challenge to an account of the determination of such values by incentive learning. Second, shifts between different motivational states control instrumental performance by a Pavlovian process that is *directly* modulated by the current

motivational state of the agent and does not depend upon incentive learning. In this respect, it is significant that Hovancik (1978) has also reported a direct effect of motivational shifts on Pavlovian-instrumental transfer, but in this case the shifts were between different levels of the same motivational state rather than between states. He gave rats either discriminative leverpress training or noncontingent Pavlovian stimulus-food pairings under different levels of food deprivation in an initial training phase before testing instrumental performance to the Pavlovian and discriminative stimuli following a shift in the deprivation level. Whereas performance to the discriminative stimulus reflected the deprivational level during both the initial training and test phases, control by the Pavlovian stimulus was determined solely by the degree of hunger at the time of testing in the second phase.

INCENTIVE LEARNING REVISITED

This Pavlovian analysis, of course, does not imply that the value of an outcome as an instrumental goal cannot be changed by a shift between motivational states, but such a revaluation, according to incentive learning theory, should depend upon animals' receiving prior exposure to the outcomes in the test state so that they can learn about their relative values in this state. Dickinson and Dawson (1988, 1989) examined this prediction in a series of studies. Of course, any assessment of the motivational control of outcome value by motivational transitions must control the Pavlovian influences. Dickinson and Dawson did this by using a concurrent training procedure in which one action, leverpressing or chain pulling by hungry rats, produced one outcome, the sucrose solution or food pellets, and the other action produced the alternative outcome. As we have pointed out, such a concurrent training procedure ensures that the contextual cues are equally associated with both the sucrose solution and the food pellets, thus equating the Pavlovian influence of these cues at the time of testing for all animals. This test then examined the propensity of the animals to press the lever in extinction while they were thirsty.

In addition, all the animals also received noncontingent exposure to the outcomes in the absence of the manipula at some point prior to testing. During this exposure period, when the motivational state of the animals was alternated between hunger and thirst, the experimental group was presented with the sucrose solution and food pellets on days when they were thirsty, thus allowing them to learn about the relative values of the two outcomes in the test state. If this incentive learning experience brings about control over the relative outcome values by the animals' current motivational state, we should expect those trained to leverpress for the sucrose to press more than those for which the food pellet was the outcome of this action. This is just what Dickinson and Dawson (1988, 1989) observed. Moreover, this irrelevant-incentive effect, which in this case must have been mediated by the instrumental contingency, occurred when the incentive

learning treatment was given prior to instrumental training, in the middle of this training, or between training and testing. Again, this finding is just what would be predicted by the incentive learning account. As long as the animals have the opportunity to assign outcome values to the sucrose solution and food pellets that are appropriate to the state of thirst prior to the test, we should expect to find the difference in the relative values to be reflected in instrumental performance.

Of course, this interpretation of the irrelevant-incentive effect assumes that the critical feature of the exposure procedure is the presentation of the outcomes under thirst. This appears to be so; across a series of experiments, Dickinson and Dawson (1988, 1989) failed to detect an effect of the type of outcome used to train leverpressing in control groups that received the outcomes on days when the animals were hungry during the exposure phase.

In conclusion, incentive learning appears to play a role in the control of instrumental actions across a transition between motivational states that is comparable to its role in mediating behavioral adjustments across shifts in level of a single motivational state once, of course, the influence of Pavlovian processes is controlled. A shift between two different motivational states will bring about an appropriate change of outcome value (and hence a change in instrumental performance mediated by knowledge of the action-outcome contingency) only if the animals have previously had the opportunity to sample the outcomes in the shifted state.

We should conclude, however, by noting an apparent exception to this claim, which comes from studies of transitions from water to food deprivation. As we have already noted, following training under thirst, hungry rats press more if they have been trained with the sucrose solution rather than saline as the outcome (Dickinson & Balleine, 1990). This difference presumably reflects the fact that the sucrose solution is relevant to the test state of hunger, whereas the saline is not. In contrast to hunger-to-thirst and thirst-to-sodium appetite transitions, this irrelevant-incentive effect is observed in the absence of any opportunity for incentive learning under the test state of hunger, and even after concurrent training that should control the influence of the Pavlovian motivational process. In these studies, thirsty rats were trained to press for either a sucrose solution or saline and to chain pull for the other outcome before being tested for leverpressing in extinction when hungry (Dickinson & Balleine, 1990). Uniformly, we found that the animals pressed more on test when the sucrose solution rather than the saline had been the training outcome. In other words, a shift to the state of hunger appeared to endow the sucrose solution with a higher value than the saline, even though the animals had never been previously exposed to these outcomes when food deprived.

In order to accommodate this finding within incentive learning theory, we suggested that water-deprived rats may in fact be hungry as well as thirsty. Although it is well known that depriving a rat of either water or food

reduces consumption of the other commodity during the deprivation period (e.g., Verplanck & Hayes, 1953), the cause of the reduction appears to be different in the two cases (Toates, 1978). In the case of water deprivation, there appears to be a central inhibition of feeding behavior (e.g., Rolls & McFarland, 1973), including food-reinforced instrumental actions in the case of the rat (Van Hemel & Myers, 1970), which results in the accumulation of a nutritional deficit across periods of water deprivation. Presumably, this inhibition functions to defend the water balance of the animal against the challenge produced by the ingestion of nutrients. Thus, water-deprived rats may well be hungry as well as thirsty, although the behavioral manifestation of the hunger is inhibited. However, this inhibition need not prevent the water-deprived animal from learning about the relative values of commodities, such as the sucrose solution and the saline, with respect to the state of hunger. This means that the incentive learning that is necessary for the instrumental control of performance under hunger could well take place during concurrent instrumental training under water deprivation.

The situation is very different in the case of a transition from hunger to thirst, for there is no reason to believe that a food-deprived rat is in any sense thirsty. A reduction in food intake should not present a challenge to fluid balance and, in fact, rats tend to drink more water than they require under food deprivation (Morrison et al., 1967). Thus, the food-deprived animals have no opportunity to learn about the relative values of the sucrose solution and food pellets with respect to the state of thirst during concurrent training under food deprivation. In order for the state of thirst to gain control over the relative outcome values assigned to these two commodities, it should be necessary for the animal to be exposed to them while thirsty. In accord with this prediction, we have extensively documented the necessity of incentive learning for the motivational control of instrumental performance across a hunger-to-thirst transition.

DUAL-PROCESS THEORY OF INSTRUMENTAL MOTIVATION

In this paper, we have discussed the motivational control of instrumental, goal-directed actions. By goal directed, we mean any activity whose acquisition depends upon the instrumental contingency between the action and the outcome (instrumental criterion) and whose performance is sensitive to the current value of the outcome (goal criterion). Our analysis has revealed that such behavior is controlled by two motivational processes.

Pavlovian Motivational Process

In accord with traditional two-process theory (Rescorla & Solomon, 1967; Trapold & Overmier, 1972), contextual and discriminative stimuli acquire general motivational and/or cuing properties through their Pavlovian association with the outcomes during acquisition of the

instrumental action. This motivational influence appears to have two main properties. First, the modulating effect of such stimuli depends upon the relevance of the associated outcome to the current motivational state of the agent and, second, this effect of the motivational state is direct and appears to operate even if the agent has not previously experienced the outcome in that state.

Incentive Learning Process

The second motivational process operates through the agent's knowledge of the action-outcome relation, and we have followed Tolman in assuming that motivational variables act by determining and controlling the value assigned to the outcome. It is clear, however, that the assignment is not automatic or direct but, in contrast to the Pavlovian process, depends upon the agent's past experience with the outcome in the relevant motivational state. Tolman (1949a, 1949b) argued as much when he proposed that the current value of an outcome reflects the degree to which a motivational state is capable of exciting or activating a representation of this event and that the connection mediating this excitation may well have to be acquired through experience with the outcome in that state.

Although Tolman's cathexis theory anticipates that an increase in motivational level will not automatically produce an elevation of performance, it predicts that a reduction in motivational level should result in an immediate corresponding decrement in instrumental performance. A review of the relevant literature reveals, however, that such performance typically persists at a level appropriate to the motivated state. In order to induce sensitivity not only to upshifts in motivational level but also to downshifts, it is necessary for the agent to have previously experienced the outcome of the instrumental action in the test state.

It was for this reason that we have appealed to the more general concept of incentive learning. Incentive learning refers to the assignment of a value to an instrumental outcome through direct exposure to the outcome. Primary motivational states interact with this incentive learning process in two ways. The first is by *determining* the agent's reactions to the outcome at the time of contact. Thus, exposure to a nutritious and/or palatable commodity will lead to the assignment of a higher value when the agent is hungry rather than sated. However, this assigned value, although determined by the current motivational state at the time of exposure, is not automatically *controlled* by or conditional upon this motivational state; in other words, a value assigned under one level of motivation appears to generalize across shifts in motivational level, and possibly also across shifts in state. In order for outcome value to change appropriately across shifts in motivational level or state, it appears to be necessary that the agent have prior experience with the outcome in the two states. To what extent such incentive discrimination training will endow other, novel motivational levels or states with control over outcome value is as yet unknown.

In summary, our claim is that the motivational control over goal-directed actions is, at least in part, indirect and mediated by learning about one's own reactions to primary incentives. By this process of incentive learning, goal-directed actions are liberated from the tyranny of primary motivation. It is a feature of our own goal-directed behavior that it can be deployed in the pursuit of abstract goals—honor, beauty, truth, and prestige, for example, the value of which must surely be socially conditioned and learned. As we have seen, however, the role of learning in establishing value is not peculiar to such goals for it is a feature of all desires, even those related to the most basic and primary motivational states.

REFERENCES

- ADAMS, C. D., & DICKINSON, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **33B**, 109-122.
- ASRATYAN, E. A. (1974). Conditioned reflex theory and motivational behavior. *Acta Neurobiologiae Experimentalis*, **34**, 15-31.
- BAEYENS, F., EELEN, P., VAN DEN BERG, H., & CROMBEZ, G. (1990). Flavor-flavor and color-flavor conditioning in humans. *Learning & Motivation*, **21**, 434-455.
- 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., & DICKINSON, A. (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *Quarterly Journal of Experimental Psychology*, **43B**, 279-296.
- BALLEINE, B., & DICKINSON, A. (1992). Signalling and incentive processes in instrumental reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, **45B**, 285-301.
- BERRIDGE, K. C. (1991). Modulation of taste affect by hunger, caloric satiety, and sensory-specific satiety in the rat. *Appetite*, **16**, 103-120.
- BERRIDGE, K. C., FLYNN, F. W., SCHULKIN, J., & GRILL, H. J. (1984). Sodium depletion enhances salt palatability in rats. *Behavioral Neuroscience*, **98**, 652-660.
- BERRIDGE, K. C., GRILL, H. J., & NORGREN, R. (1981). Relation of consummatory responses and preabsorptive insulin release to palatability and learned taste aversions. *Journal of Comparative & Physiological Psychology*, **95**, 363-382.
- BINDRA, D. (1972). A unified account of classical conditioning and operant training. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 453-482). New York: Appleton-Century-Crofts.
- BINDRA, D. (1978). How adaptive behavior is produced: A perceptual motivational alternative to response-reinforcement. *Behavior & Brain Sciences*, **1**, 41-52.
- BOLLES, R. C. (1972). Reinforcement, expectancy, and learning. *Psychological Review*, **95**, 394-409.
- BOLLES, R. C. (1975). *Theory of motivation*. New York: Harper & Row.
- BOLLES, R. C., HOLTZ, R., DUNN, T., & HILL, W. (1980). Comparison of stimulus learning and response learning in a punishment situation. *Learning & Motivation*, **11**, 78-96.
- CABANAC, M. (1971). Physiological role of pleasure. *Science*, **173**, 1103-1107.
- CAMPBELL, D. H., CAPALDI, E. D., SHEFFER, J. D., & BRADFORD, J. P. (1988). An examination of the relationship between expectancy learning and preference conditioning. *Learning & Motivation*, **19**, 162-182.
- 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., & MYERS, D. E. (1982). Taste preferences as a function of food deprivation during original taste exposure. *Animal Learning & Behavior*, **10**, 211-219.

- CAPALDI, E. D., MYERS, D. E., CAMPBELL, D. H., & SHEFFER, J. D. (1983). Conditioned flavor preferences based on hunger level during original flavor exposure. *Animal Learning & Behavior*, **11**, 107-115.
- CAPALDI, E. D., SHEFFER, J., & OWENS, J. (1991). Food deprivation and conditioned flavor preferences based on sweetened and unsweetened foods. *Animal Learning & Behavior*, **19**, 361-368.
- CAPALDI, E. D., VIVEIROS, D. M., & DAVIDSON, T. L. (1981). Deprivation stimulus intensity and incentive factors in the control of instrumental performance. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 140-149.
- COLWILL, R. C., & RESCORLA, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 120-132.
- COLWILL, R. C., & RESCORLA, R. A. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 55-104). New York: Academic Press.
- CROCETTI, C. P. (1962). Drive level and response strength in the bar-pressing apparatus. *Psychological Reports*, **10**, 563-575.
- DAVIDSON, T. L. (1987). Learning about deprivation intensity stimuli. *Behavioral Neuroscience*, **101**, 198-208.
- DAVIDSON, T. L., FLYNN, F. W., & JARRARD, L. E. (1992). Potency of food deprivation intensity cues as discriminative stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 174-181.
- DICKINSON, A. (1980). *Contemporary animal learning theory*. Cambridge: Cambridge University Press.
- DICKINSON, A. (1986). Re-examination of the role of the instrumental contingency in the sodium-appetite irrelevant incentive effect. *Quarterly Journal of Experimental Psychology*, **38B**, 161-172.
- DICKINSON, A. (1989). Expectancy theory in animal conditioning. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 279-308). Hillsdale, NJ: Erlbaum.
- DICKINSON, A., & BALLEINE, B. (1990). Motivational control of instrumental performance following a shift from hunger to thirst. *Quarterly Journal of Experimental Psychology*, **42B**, 413-431.
- DICKINSON, A., & DAWSON, G. R. (1987a). Pavlovian processes in the motivational control of instrumental performance. *Quarterly Journal of Experimental Psychology*, **38B**, 201-213.
- DICKINSON, A., & DAWSON, G. R. (1987b). The role of the instrumental contingency in the motivational control of performance. *Quarterly Journal of Experimental Psychology*, **39B**, 77-93.
- 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 Journal of Experimental Psychology*, **41B**, 99-112.
- DICKINSON, A., & NICHOLAS, D. J. (1983). Irrelevant incentive learning during instrumental conditioning: The role of drive-reinforcer and response-reinforcer relationships. *Quarterly Journal of Experimental Psychology*, **35B**, 249-263.
- EISMAN, E., THEIOS, J., & LINTON, M. (1961). Habit strength as a function of drive in a bar pressing situation. *Psychological Reports*, **9**, 583-590.
- ESTES, W. K. (1943). Discriminative conditioning: 1. A discriminative property of conditioned anticipation. *Journal of Experimental Psychology*, **32**, 150-155.
- EVERITT, B. J., & STACEY, P. (1987). Studies of instrumental behavior with sexual reinforcement in male rats (*Rattus norvegicus*): 2. Effects of preoptic area lesions, castration, and testosterone. *Journal of Comparative Psychology*, **101**, 407-419.
- FITZSIMONS, J. T. (1961). Drinking by nephrectomized rats injected with various substances. *Journal of Physiology*, **155**, 563-579.
- 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, NJ: Erlbaum.
- GORMEZANO, I., & TAIT, R. W. (1976). The Pavlovian analysis of instrumental conditioning. *Pavlovian Journal of Biological Sciences*, **11**, 37-55.
- GRICE, G. R., & DAVIS, J. D. (1957). Effect of irrelevant thirst motivation on a response learned with food reward. *Journal of Experimental Psychology*, **53**, 347-352.
- GRILL, H. J., & BERRIDGE, K. C. (1985). Taste reactivity as a measure of the neural control of palatability. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in psychobiology and physiological psychology* (Vol. 11, pp. 1-61). Orlando, FL: Academic Press.
- GRINDLEY, G. C. (1932). The formation of a simple habit in guinea pigs. *British Journal of Psychology*, **23**, 127-147.
- HEATHERS, G. L., & ARAKELIAN, P. (1941). The relationship between strength of drive and rate of extinction of a bar-pressing reaction in the rat. *Journal of General Psychology*, **24**, 243-258.
- HENDERSEN, R. W., & GRAHAM, J. (1979). Avoidance of heat by rats: Effects of thermal context on the rapidity of extinction. *Learning & Motivation*, **10**, 351-363.
- HERSHBERGER, W. A. (1986). An approach through the looking-glass. *Animal Learning & Behavior*, **14**, 443-451.
- HORENSTEIN, B. R. (1951). Performance of conditioned responses as a function of strength of hunger drive. *Journal of Comparative & Physiological Psychology*, **44**, 210-224.
- HOVANCIK, J. R. (1978). The effect of deprivation level during non-contingent pairings and instrumental learning on subsequent instrumental performance. *Learning & Motivation*, **9**, 1-15.
- HULL, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- IRWIN, F. W. (1971). *Intentional behavior and motivation: A cognitive theory*. Philadelphia: Lippincott.
- JENKINS, H. M., & HARRISON, R. H. (1960). The effects of discrimination training on auditory generalization. *Journal of Experimental Psychology*, **59**, 246-253.
- KING, D. L. (1979). *Conditioning: An image approach*. New York: Gardner.
- KOCH, S., & DANIEL, W. J. (1945). The effect of satiation on behavior mediated by a habit of maximum strength. *Journal of Experimental Psychology*, **35**, 167-187.
- KONORSKI, J. (1967). *Integrative activity of the brain: An interdisciplinary approach*. Chicago: University of Chicago Press.
- KRIECKHAUS, E. E., & WOLF, G. (1968). Acquisition of sodium by rats: Interaction of innate mechanisms and latent learning. *Journal of Comparative & Physiological Psychology*, **65**, 197-201.
- LOPEZ, M., BALLEINE, B., & DICKINSON, A. (1992a). Incentive learning and the motivational control of instrumental performance by thirst. *Animal Learning & Behavior*, **20**, 322-328.
- LOPEZ, M., BALLEINE, B., & DICKINSON, A. (1992b). Incentive learning following reinforcer devaluation is not conditional upon motivational state during re-exposure. *Quarterly Journal of Experimental Psychology*, **45B**, 265-284.
- MEACHUM, C. L. (1990). The role of response-contingent incentives in lithium chloride-mediated suppression of an operant response. *Quarterly Journal of Experimental Psychology*, **42B**, 175-195.
- MILLER, N. E. (1963). Some reflections on the law of effect produce a new alternative to drive reduction. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 65-112). Lincoln: University of Nebraska Press.
- MILLER, S., & KONORSKI, J. (1969). On a particular form of conditioned reflex. *Journal of Experimental Analysis of Behavior*, **12**, 187-189. [Translation by B. F. Skinner of: Sur une forme particulière des reflex conditionnels. *Les Comptes Rendus des Seances de la Société Polonaise de Biologie*, 1928, **49**, 1155-1157].
- MORGAN, M. J. (1974). Resistance to satiation. *Animal Behaviour*, **22**, 449-466.
- MORRISON, S. D., MACKAY, C., HURLBRINK, E., WIER, J. K., NICK, M. S., & MILLER, F. K. (1967). The water exchange and polyuria of rats deprived of food. *Quarterly Journal of Experimental Physiology*, **52**, 51-67.
- MOWRER, O. H. (1960). *Learning theory and behavior*. New York: Wiley.
- PAVLOV, I. P. (1927). *Conditioned reflexes*. Oxford: Oxford University Press.
- PAVLOV, I. P. (1932). The reply of a physiologist to a psychologist. *Psychological Review*, **39**, 91-127.
- PELCHAT, M. L., GRILL, H. J., ROZIN, P., & JACOBS, J. (1983). Qual-

- ity of acquired responses to taste by *Rattus norvegicus* depends upon type of associated discomfort. *Journal of Comparative & Physiological Psychology*, **97**, 140-153.
- PERIN, C. T. (1942). Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *Journal of Experimental Psychology*, **30**, 93-113.
- PETERSON, D., & MCHOSE, J. H. (1980). The residual deprivation effect in instrumental conditioning as a function of deprivation level at the time of exposure to the reinforcer. *Animal Learning & Behavior*, **8**, 642-646.
- 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., & SOLOMON, R. L. (1967). Two-process learning theory: Relationship between Pavlovian conditioning and instrumental learning. *Psychological Review*, **74**, 151-182.
- REVUSKY, S. H. (1967). Hunger level during food consumption: Effects on subsequent preference. *Psychonomic Science*, **7**, 109-110.
- REVUSKY, S. H. (1968). Effects of thirst level during consumption of flavored water on subsequent preference. *Journal of Comparative & Physiological Psychology*, **66**, 777-779.
- ROLLS, B. J., & MCFARLAND, D. J. (1973). Hydration releases inhibition of feeding produced by intracranial angiotensin. *Physiology & Behavior*, **11**, 881-884.
- ROZIN, P., & FALLON, A. E. (1987). A perspective on disgust. *Psychological Review*, **94**, 23-41.
- SACKETT, R. S. (1939). The effect of strength of drive at the time of extinction upon resistance to extinction in rats. *Journal of Comparative Psychology*, **27**, 411-431.
- SALTZMAN, I., & KOCH, S. (1948). The effect of low intensities of hunger on the behavior mediated by a habit of maximum strength. *Journal of Experimental Psychology*, **38**, 347-370.
- SHEFFIELD, F. D. (1965). Relation between classical conditioning and instrumental learning. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium* (pp. 302-322). New York: Appleton-Century-Crofts.
- SKINNER, B. F. (1932). On the rate of formation of a conditioned reflex. *Journal of General Psychology*, **7**, 274-285.
- SKINNER, B. F. (1936). Conditioning and extinction and their relation to drive. *Journal of General Psychology*, **14**, 296-317.
- SPENCE, K. W. (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press.
- SUTTON, R. S., & BARTO, A. G. (1981). An adaptive network that constructs and uses an internal model of its world. *Cognition & Brain Theory*, **4**, 217-246.
- THORNDIKE, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.
- TOATES, F. M. (1978). A physiological control theory of hunger-thirst interaction. In D. A. Booth (Ed.), *Hunger model* (pp. 347-374). London: Academic Press.
- TOLMAN, E. C. (1949a). The nature and function of wants. *Psychological Review*, **56**, 357-369.
- TOLMAN, E. C. (1949b). There is more than one kind of learning. *Psychological Review*, **56**, 144-155.
- 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.
- TRAPOLD, M. A., & OVERMIER, J. B. (1972). The second learning process in instrumental learning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 427-452). New York: Appleton-Century-Crofts.
- VAN HEMEL, P. E., & MYERS, J. S. (1970). Control of food-motivated instrumental behavior in water-deprived rats by prior water and saline drinking. *Learning & Motivation*, **1**, 86-94.
- VERPLANCK, W. S., & HAYES, J. R. (1953). Eating and drinking as a function of maintenance schedule. *Journal of Comparative & Physiological Psychology*, **46**, 327-333.
- WEBB, W. B. (1955). Drive stimuli as cues. *Psychological Reports*, **1**, 287-298.
- WETZEL, R. J. (1959). The effect of experience with a taste reward. *Journal of Comparative & Physiological Psychology*, **52**, 267-271.
- WOODS, S. C., VASSELLI, J. R., & MILAN, K. M. (1977). Iron appetite and latent learning in rats. *Physiology & Behavior*, **19**, 623-626.
- YOUNG, P. T. (1961). *Motivation and emotion*. New York: Wiley.

NOTES

1. This analysis assumes that stimuli are defined allocentrically rather than egocentrically. If one is prepared to accept a behaviorally contingent definition of stimuli, it is possible to give an account of Grindley's (1932) observation in terms of stimulus-outcome relations. Thus, it could be argued that the instrumental contingency between head turning, say to the left, and the outcome established a stimulus-outcome relation between the contextual stimuli on the left and the carrot, which would result in the animal's orienting to that side. Bindra (1972, 1978), for example, has offered such an account of instrumental conditioning. As the definition of such stimuli is conditional upon an action, however, sensitivity to stimulus-outcome and action-outcome relations are indistinguishable at a behavioral level, and will not be considered further. In the present discussion, we shall assume that stimuli are defined independently of behavior.

2. This characterization of instrumental learning follows Pavlov and his students (Asratyan, 1974; Pavlov, 1932) in assuming that the instrumental performance is mediated by a connection from the outcome unit to the response unit. Although there are major problems with this so-called bidirectional theory (see Dickinson, 1980; Gormezano & Tait, 1976), we employ this account in the present context for its explanatory simplicity.

3. This model assumes that flavor-aversion conditioning does not establish the flavor as an aversive instrumental outcome, but rather simply attenuates its attractiveness. Consequently, the model cannot explain Meachum's (1990) demonstration that such a flavor will act as an instrumental punisher.

4. All the irrelevant-incentive studies by Dickinson and his colleagues employed a relatively concentrated 20% sucrose solution in an attempt to minimize any generalization of incentive learning from the animals' experience with water in their home cages. In the present context, it is important to note that even hypertonic sucrose taken in relatively small quantities by mouth should not act as a thirst stimulus. Sucrose is hydrolyzed by the digestive enzymes into its constituent monosaccharides, which have minimal impact on plasma osmolality and do not act as a thirst stimulus (e.g., Fitzsimons, 1961).

(Manuscript received October 27, 1992;
revision accepted for publication May 21, 1993.)