Neural bases of food-seeking: Affect, arousal and reward in corticostriatolimbic circuits

Bernard W. Balleine*

Department of Psychology and the Brain Research Institute, University of California, Box 951563, Los Angeles, CA 90095-1563, United States

Received 28 July 2005; accepted 25 August 2005

Abstract

Recent studies suggest that there are multiple ‘reward’ or ‘reward-like’ systems that control food seeking; evidence points to two distinct learning processes and four modulatory processes that contribute to the performance of food-related instrumental actions. The learning processes subserve the acquisition of goal-directed and habitual actions and involve the dorsomedial and dorsolateral striatum, respectively. Access to food can function both to reinforce habits and as a reward or goal for actions. Encoding and retrieving the value of a goal appears to be mediated by distinct processes that, contrary to the somatic marker hypothesis, do not appear to depend on a common mechanism but on emotional and more abstract evaluative processes, respectively. The anticipation of reward on the basis of environmental events exerts a further modulatory influence on food seeking that can be dissociated from that of reward itself; earning a reward and anticipating a reward appear to be distinct processes and have been doubly dissociated at the level of the nucleus accumbens. Furthermore, the excitatory influence of reward-related cues can be both quite specific, based on the identity of the reward anticipated, or more general based on its motivational significance. The influence of these two processes on instrumental actions has also been doubly dissociated at the level of the amygdala. Although the complexity of food seeking provides a hurdle for the treatment of eating disorders, the suggestion that these apparently disparate determinants are functionally integrated within larger neural systems may provide novel approaches to these problems.

D 2005 Elsevier Inc. All rights reserved.

Keywords: Goal-directed action; Habit learning; Instrumental conditioning; Pavlovian conditioning; Incentive learning; Motivation; Striatum; Thalamus; Prefrontal cortex; Amygdala

1. Introduction

There has been a recent trend towards identifying the processes involved in obesity with those associated with addictive behavior generally and with drug addiction in particular. For example, in a recent series of papers, Volkow and colleagues have established that binding at the dopamine D2 receptor in obese subjects, i.e., those with a body mass index over 30, is reduced in similar fashion to that of individuals addicted to drugs of abuse [119–122]. A feature of these, and similar [24], accounts is that, often in the interests of a simple story, they focus on one factor, brain dopamine, as the causal factor, not just in pathological food intake but in its sequela, notably in food seeking or pursuit. The operation of the reward system is commonly argued to link intake and pursuit and, indeed, since the discovery of self-stimulation, students of neuroscience have felt strongly predisposed to the view that there is a central reward system in the brain, that it is monolithic and that it involves midbrain dopaminergic neurons and particularly their projection via the medial forebrain bundle to limbic structures in the ventral forebrain [61,87,137]. It has appeared, therefore, to be a reasonable leap to propose that pathologies of brain dopamine are associated, more or less directly, with pathologies of the ‘reward system’ and so with pathological food seeking [44]. Indeed, evidence that, in addition to reduced D2 receptor binding, drug addicts have increased genetic variation associated with the D2 receptor has raised the specter of a ‘reward gene’ [25,26]. Of course, it is equally possible that this evidence points to a corollary of addiction rather than its efficient cause. But these issues aside, the real problem with this approach is that it over-simplifies our
understanding of the complex nature of the processes that contribute to both normal and abnormal food seeking. A number of recent papers have, as a consequence, unnecessarily conflated the processes that contribute to the compulsive pursuit of food with those that control goal-directed actions [63, 66, 85] and still further with those that control responses elicited by stimuli associated with food [74]. Although the operation of these processes objectively affects the rate of food seeking, recent evidence suggests that they each have distinct determinants. This review will attempt to tease these various influences apart with reference to recent research that has identified not one but potentially five ‘reward’ or ‘reward-related’ processes in the brain; that is to say, five systems that function to influence food seeking either directly, through learning, or indirectly, by modifying performance.

2. Reward and reinforcement

The recent literature concerning drug seeking in addicts has focused attention on the compulsive or habitual nature of these responses revealed particularly in their persistence, even in the face of sometimes quite extreme negative consequences, and their sensitivity to drug-related cues, an observation that has informed various theories of relapse [29, 54, 95, 107]. Many of the ideas that have been expressed in these recent papers have their root in now classical theories of habit learning, associated most notably with Hull [76], that explain the acquisition of actions instrumental to gaining access to rewarding events in terms of the operation of a stimulus–response/reinforcement (S–R) architecture. From this perspective, addictive drugs reinforce or strengthen associations between continguously active sensory and motor processes allowing the sensory process subsequently to elicit the motor response in a manner that is no longer regulated by its consequences.

Although it is a straightforward matter to apply these ideas to drug addiction, it is much less clear whether and to what degree they apply directly to activities associated with natural rewards like food. Although S–R theorists regarded food seeking, like compulsive drug seeking, as a form of habit, what evidence there is for this claim has really only emerged relatively recently in studies assessing the effect of post-training reinforcer devaluation on instrumental performance. For example, Holman [75] was able to show that lever press responses in thirsty rats reinforced on an interval schedule by access to a saccharin solution were maintained in extinction even after the saccharin had been devalued by pairing its consumption with illness. It is important to recognize how maladaptive the lever pressing was in Holman’s rats. Although the pairing with illness resulted in the rats no longer consuming or even contacting the previously palatable (but now poisonous) saccharin, their subsequent extinction performance on the lever continued at a rate comparable to that of rats for which the saccharin was not devalued.

Several years later in a replication of Holman’s experiment, Adams and Dickinson [1] found, in contrast, that, when lever pressing in hungry rats was reinforced either continuously or on a ratio schedule by sugar pellets, devaluation of the pellets strongly attenuated subsequent performance on the lever. Although several features of the two studies differed, Dickinson, Nicholas and Adams [51] later showed that interval schedules of reinforcement were particularly apt to produce habitual responses; i.e., responses that are no longer dependent on the current value of their consequences; when previously reinforced by sugar on a ratio schedule lever pressing was sensitive to devaluation whereas when reinforced on an interval schedule it was not. These findings provide direct evidence that, over and above a habit or S–R process, the performance of instrumental actions can also be goal-directed; it can reflect encoding of the relationship between action and outcome. Furthermore, they show that both processes can be engaged depending on the relationship between instrumental performance and reward delivery. When reward delivery is constrained by time so that changes in the rate of performance have little if any effect on the rate of reward, actions tend to become habitual. When rate of reward is proportional to the rate of performance, however, actions tend to be goal-directed. It is also worth noting that the fact that the same event, sucrose in this case, could serve both as the goal of an action and to reinforce S–R associations must raise immediate questions regarding the notion of single or monolithic ‘reward’ system responsible for all changes in instrumental performance.

Recent experimentation has only made these questions more pointed. For example, several recent studies have found evidence that damage to the lateral region of the dorsal striatum (DLS) renders rats incapable of developing simple S–R solutions to various maze discrimination problems suggesting that this region may be important in the formation of associations of this kind [46, 92, 99]. Anatomically, the DLS appears to be well suited to this functional role, maintaining strong connections with sensorimotor regions of the neocortex [93]. Furthermore, this region receives a dense projection from the midbrain dopaminergic neurons that electrophysiological studies suggest may play a reinforcing role, modulating plasticity of the basal ganglia to the cortex [105]. In a recent study we attempted to provide more direct evidence for the involvement of the DLS in habit learning by assessing the effect of cell-body lesions of this area on the acquisition and performance of instrumental actions trained on an interval schedule of reinforcement as well as on sensitivity of performance to the devaluation of the instrumental outcome [134]. The strong view that the DLS mediates S–R learning predicts that acquisition and subsequent performance of actions reinforced on interval schedules should be severely attenuated. Against this prediction we found that acquisition was normal and subsequent performance was only moderately affected by the lesion. The most striking effect was, however, the change in the influence of outcome devaluation. Whereas the instrumental performance of sham-lesioned controls showed no sensitivity whatever to outcome devaluation by conditioned taste aversion, replicating previous findings, the DLS-lesioned group showed clear sensitivity to this treatment [134]. This result was specific to the DLS; lesions of the dorsomesial striatum did not increase sensitivity to outcome devaluation. The lesions of the DLS, therefore, effectively abolished...
habitual responding and rendered instrumental actions goal-directed.

This result suggests that both habit and goal-directed learning processes are concurrently engaged but that one or other process predominates depending upon the circumstances during training. It appears, therefore, that food seeking can, at the very least, be accomplished through two distinct means; either by habitual or compulsive performance of responses previously reinforced by access to food or by more deliberate, goal-directed actions aimed at achieving access to specific rewarding events. In recent years much of the work in my lab has been focused on developing an understanding of the behavioral and neural bases of this latter class of activity and the remainder of this paper will be concerned with the processes that influence its acquisition and performance.

3. Goal-directed learning

3.1. Behavioral considerations

Instrumental conditioning in rodents provides a very accurate model of goal-directed action in humans. Not only are rodent actions sensitive to changes in the value of the goal or outcome with which they are associated but they are also highly sensitive to changes in their causal consequences; rats will stop responding if performance no longer delivers the instrumental outcome and will stop responding even faster if their responding cancels an otherwise freely available food [45,53]. Hammond [68] developed a schedule with which he could manipulate independently the probability of an outcome (water for thirsty rats) given performance of a particular response (lever pressing; i.e., \( p(O/R) \)) and the probability of an outcome in the absence of a response (\( p(O/noR) \)). He found that performance was reduced as the probability of a non-contiguous outcome was increased despite the fact that contiguity (i.e., \( p(O/R) \)) was kept constant and at a rate that ordinarily maintained substantial levels of performance.

A number of studies have confirmed this observation and extended it to establish that rats are sensitive to the selective degradation of one action–outcome contingency in a situation where another contingency is maintained intact [13]. In one study, hungry rats were trained to perform two actions, lever pressing and chain pulling, with one action earning food pellets and the other a polycope solution. Both actions were trained on Hammond’s schedule with \( p(O/R) \) set at 0.05. After training, one contingency was degraded such that, in addition to being earned by performing its associated action, one of the outcomes was also delivered non-contiguously at the same probability but in each second without a response, i.e., \( p(O/noR) = 0.05 \). Thus, the experienced probability of the delivery of that particular outcome in any one second was the same whether the animals performed that action or not, ensuring that one and not the other instrumental contingency was degraded. If instrumental performance reflects the rats’ sensitivity to the contingent relation between the performance of an action and its specific consequences then degrading the specific action–outcome contingency in this way should result in reduction in the performance of that specific action relative to performance of the other action. Again, this is exactly what was found; only performance of the action the outcome of which was the same as that delivered non-contingently was reduced providing evidence that performance was sensitive to the action–outcome contingency [13,31,50,126].

Although this contingency framework provides a good first approximation of the learning rules that mediate the encoding of act–goal associations (i.e., action–outcome or A–O learning), it cannot be the whole story; simple reflection on the differing effects of training on ratio and interval schedules of reinforcement suggests that much. Rather, and as implied above, it appears that instrumental learning reflects the correlation between the rate of performance of a particular action and the rate of delivery of its specific outcome calculated individually for that action through time [17,47]. Psychologically, this learning should clearly be regarded as declarative; performance reflects the ability of animals to utilize information about the action–outcome relationship in the face of changing expectations of reward [127]. Nevertheless, despite arguments regarding the function of the hippocampus in declarative learning of this kind [57,113,114], in several series of studies we were unable to find any clear evidence for the involvement of the hippocampus or its outflow through anterior thalamus in instrumental learning [34,39,40]. These early experiments did, however, find evidence for the involvement of the mediodorsal thalamus as well as one of its main cortical projection areas – the prelimbic region of the medial prefrontal cortex (PL) – in this form of learning. Unlike the hippocampus, cell body lesions of these areas were effective in abolishing rats’ sensitivity to both outcome devaluation and to selective degradation of the instrumental contingency [13,36,39]. In the context of the effects of DLS lesions mentioned above it is interesting to note that these lesions appeared to render the rats’ instrumental performance habitual.

More recent evidence suggests, however, that the involvement of the prefrontal cortex in goal-directed learning is time-limited. In a recent series we found clear evidence that only damage to the PL made prior to instrumental training had any effect on conditioning; lesions made after training was complete had no effect on outcome devaluation [138]. This suggested to us that, although the PL was clearly involved in goal-directed learning it was not the locus of encoding the action–outcome association. Rather, on the basis of its strong connections with sensory, visceral and emotional areas, we felt it likely that the PL conveys the rate of goal or outcome delivery that is then associated with information regarding the action and its rate of performance in some distal, effenter structure. The PL has two well-documented efferents; one arising predominantly in layer II and projecting to the core of the nucleus accumbens [56] and a second arising predominantly in layers V/VI and projecting to the dorsomedial or associative striatum (DMS) [18]. For reasons documented below, the results of other work have led us to believe that the former plays an important role in instrumental performance but not in instrumental learning.
3.2. Action–outcome encoding in the dorsal striatum

In fact, the DMS is an excellent candidate for the locus of instrumental learning. It is a critical component in the associative cortico-basal ganglia circuit and receives inputs from association cortices such as the PL as well as the premotor or medial agranular cortex involved in the action monitoring and programming implicated in executive processes [98,103], and projections from the pDMS are in a position to influence downstream motor control networks in the brainstem as well as the motor thalamocortical reentrant network [98]. The posterior part of the DMS (pDMS) also receives inputs from the basolateral amygdala [79], a structure that, according to recent evidence, reviewed briefly below, mediates the assignment of incentive value to the consequences of instrumental actions [16]. In accord with this suggestion, electrophysiological studies measuring neural activity in the associative striatum or caudate nucleus in primates, the homologue of the DMS in rats, have reported that neural activity in this region correlated with the performance of skilled movements can be modulated by the expectancy of reward [69,77]. Finally, in a recent series of experiments [136] we found direct evidence that, in contrast to manipulations of prefrontal cortex, both pre- and post-training cell-body lesions of the pDMS as well as local inactivation of this area induced by infusions of the GABA-A agonist muscimol, reduced the sensitivity of rats’ instrumental performance both to shifts in the action–outcome contingency and to post-training outcome devaluation. These manipulations again appeared to render the rats’ instrumental performance stimulus-bound and habitual [136].

The suggestion that the pDMS subserves action–outcome learning contrasts with other recent claims that the ventral [80] or the posterolateral striatum [3] mediates learning critical to the acquisition of goal-directed actions. Nevertheless, these studies only assessed changes in instrumental performance and did not directly assess changes in the content of learning. In a second recent series, therefore, we used well-established behavioral assays that unambiguously distinguish action–outcome learning from other types of learning to assess the role of the pDMS in the formation of action–outcome associations [135]. Given the evidence that NMDA receptor (NMDAR) activation is involved in long-term plasticity such as long-term potentiation in the dorsal striatum [28,89], we proposed that action–outcome encoding requires activation of NMDARs in the pDMS. This hypothesis was tested in rats that, after a period of pretraining, were given a bilateral infusion of either a selective NMDA antagonist (APV), or vehicle prior to a single learning session in which they were trained to press two levers for distinct food outcomes. The next day the rats were tested using an outcome devaluation protocol; i.e., they were allowed to consume one of the two outcomes for 1 h before a choice extinction test was given on the two levers. We found, first that APV immediately prior to training did not affect performance either during training or test but strongly attenuated the ability of the rats to use changes in outcome value to modify their instrumental performance; i.e., they appeared not to have encoded the specific action–outcome associations to which they were exposed during training. Furthermore, in subsequent experiments we found both that APV infused immediately after training did not have this effect on action–outcome encoding, nor did the infusion of APV into adjacent dorsolateral striatum [135].

3.3. The function of plasticity in the DMS

It remains entirely open at present how this plasticity functions within the larger circuit known to contribute to instrumental performance. One possibility lies in the fact that the dorsomedical striatum provides a strong input to cortical regions via a thalamocortical feedback circuit involving traditional basal ganglia circuitry [64]. The existence of parallel feedback loops arising in the cortex and coursing through striatum, midbrain, thalamus and back to the cortical origin has now been well documented [2,81] and, indeed, this description of the functional architecture of corticostriatal circuits has now largely superseded the earlier quite attractive idea of functional integration in the striatum through the convergence of diffuse cortical regions onto a discrete striatal target [82]. Nevertheless, it is entirely possible, indeed quite likely, that striatal plasticity allows functionally distinct parallel circuits to activate one another [97] to integrate functions by allowing one region of cortex to activate another via the thalamocortical feedback pathway. In this way, plasticity in the striatum could have the very important function of allowing, for example, an area of cortex involved in the representation of instrumental actions, such as medial agranular cortex, to activate a region of cortex involved in the representation of the instrumental outcome, such as the prelimbic area and vice versa. Indeed, this kind of link, when subsumed under the control of the rules that mediate striatal plasticity and placed within the appropriate corticostriatal feedback processes, could provide a sophisticated architecture capable of allowing animals both to encode and to retrieve outcomes that follow actions but also providing them with the ability to retrieve actions based on retrieved outcomes, which is a necessary component of planning and of choice and, indeed, of executive processes generally [60].

3.4. Summary: functional segregation of the dorsal striatum

Together the findings from these experimental investigations of the dorsal striatum have identified at least two distinct functional systems within adjacent regions; specifically, a circuit mediating goal-directed learning and involving the dorsomedical striatum and a circuit mediating habit or procedural learning and involving the dorsolateral striatum. Furthermore, these functions appear to be independent; damage to dorsolateral but not dorsomedical striatum has been found to render otherwise habitual actions goal-directed and damage to the dorsomedical striatum to render otherwise goal-directed actions habitual. It appears, therefore, that these two regions of the striatum, or at least distinct corticostriatal circuits involving these regions, may compete for control of instrumental performance. This functional and systems arrangement is summarized in Fig. 1.
Exactly how this competition is realized is not completely clear at the present time. Nevertheless, one possible source of competition could lie in the distinct contribution of reinforcement and reward systems to the performance of specific actions; the greater the contribution of a relatively non-specific reinforcement process to performance the less specifically goal-directed and more procedural instrumental performance appears to be. Hence, one means by which these systems could compete and through which, say, habitual processes could increasingly gain control over deliberated, goal-directed processes in food seeking, would be via increased inhibition of those sensory-specific, emotional processes that constitute the rewarding aspects of goals. As mentioned briefly above (and as is taken up in more detail below) the goal or reward value of specific foods is largely mediated by the basolateral amygdala. Recent evidence that lesions of the infralimbic cortex allow the value of the instrumental outcome to once again exert control over performance [84] is important in this context because this cortical region, via its projections onto the intercalated cells within the amygdala [19], appears able to modulate the output of the basolateral area [94]. It is possible, therefore, that these lesions have their effect by disinhibiting amygdala output, thereby increasing outcome-specific emotional processing and reducing the relative contribution of a pure reinforcement mechanism to performance. Alternatively, it is possible that lateral inhibition within the striatum itself contributes to the competition for control by these learning processes [118]. Whether long-term changes in the functioning of one or another or both of these processes are sufficient to induce aberrant food-seeking is, however, unknown at this time.

4. Reward and desire: instrumental incentives

The foregoing discussion suggests that, in instrumental conditioning, animals encode the relationship between specific actions and outcomes and are sensitive to the contingent relation between an action and goal delivery. It has long been recognized, however, that the encoding of an action–outcome association is not sufficient to determine the performance of an action. Any learning that takes the form ‘action A leads to outcome O’ can be used both to perform A and to avoid performing A. What is missing from this account is mention of the role that the value of the outcome plays in controlling instrumental performance. It is now well established that reward processes in instrumental conditioning depend critically on the ability of rats to evaluate the incentive value of the goal or outcome of its actions; i.e., the affective and motivationally relevant properties of the outcome [10,11,48,55].
Evidence for this claim can be drawn from any number of studies that have examined instrumental performance after a post-training manipulation of motivational state (e.g., thirst, sex, thermoregulation and so on; see Ref. [10] for review) although some of the best evidence has come from studies assessing the effect of shifts in food deprivation on food seeking. Post-training shifts, such as from hunger to satiety, often have very little direct effect on instrumental performance unless the effect of this shift on the incentive value of a specific nutritive event is made explicit through consummatory experience; i.e., through incentive learning [55]. With regard to instrumental responding for food, therefore, both hunger and satiety appear to act to affect performance, not because they affect drive [76], but because they affect the value of nutritive outcomes [4,8,10]. With regard to the processes controlling performance more generally, this analysis suggests that, in instrumental conditioning, rats encode both the relationship between actions and goals and the current incentive value of the goal but, more importantly, it further suggests that they integrate these sources of information to select a course of action. Indeed, it is in the evidence for the control over performance exerted by this integral that the fundamentally goal-directed quality of instrumental conditioning in rats is most forcefully revealed [12].

Considerable evidence suggests that the reward value of food is mediated by changes in taste processing. For example, specific satiety treatments have been found to be extremely effective in producing selective changes in the incentive value of instrumental outcomes and in the performance of actions that gain access to those outcomes [13,14] over and above the effects of satiety on motivation for nutrients generally or even for specific macronutrients. For example, in one study [14] hungry rats were trained to press a lever and pull a chain, with one action earning sour starch and the other salty starch, before they were sated on either the salty or sour starch and given an extinction test on the lever and chain. Although both actions earned equivalent nutrients of a similar macronutrient structure, the rats still altered their choice performance to favor the action that, in training, delivered the outcome on which they were not sated; i.e., they were able to modify their choice based on changes in taste [14].

In another series we assessed the effects of cell-body lesions of gustatory region of the insular cortex, for some time known to be involved in taste processing, although not taste detection [27], on specific-satiety induced devaluation and on incentive learning conducted after instrumental training when hungry after a shift to a sated state [15]. Although these lesions had no effect on the ability of rats to detect changes in value when they actually contacted a specific outcome on which they were sated, they were deeply amnesic when forced to choose between two actions based on their memory of satiety-induced changes in value.

These results suggest that the gustatory cortex operates as one component of an incentive system, acting to encode the taste features of the instrumental outcome as an aspect of the representation of the value of that outcome in memory. From this perspective, the gustatory cortex is not involved in detecting changes in incentive value; that would appear to require the integration of taste memory, involving the gustatory cortex, with an affective signal, apparently mediated by a different component of the incentive system [10]. Thus, changes in the value of the taste features of nutritive outcomes appear to be a function of emotional feedback; i.e., of the emotional response experienced contiguously with detection of the taste. If the emotional response is pleasant, the value of the outcome is correspondingly increased whereas if it is unpleasant it is reduced. Hence treatments that produce changes in palatability in rats, usually assessed by taste reactivity responses, are also those that most potently modify the value of the instrumental outcome [20,22]. For example, Rolls et al. [108] have provided a clear demonstration that, in humans, eating one particular food to satiety strongly reduces the pleasantness rating of that food but not other similar foods. Likewise, in rats, Berridge [21] demonstrated that, when sated on milk, ingestive taste reactivity responses were reduced and aversive taste reactivity responses were increased when milk, but not when sugar, was subsequently contacted. These kinds of data suggest, therefore, that satiety-induced changes in incentive value are not a product of general shifts in motivation but reflect variations in the association of taste features with specific emotional responses.

If incentive learning is determined by an association of sensory and emotional processes, one should suppose neural structures implicated in the formation of associations of this kind to be critically involved in this form of learning. The gustatory cortex maintains strong reciprocal connections with the amygdala [115,133] and, indeed, this connection has been implicated in taste-affect associations in a variety of paradigms [62]. The basolateral amygdala (BLA) has itself been heavily implicated in a variety of learning paradigms that have an evaluative component; for example this structure has long been thought to be critical for fear conditioning and has recently been reported to be involved in a variety of feeding related effects including sensory specific satiety [91], the control of food-related actions (see below) and in food consumption elicited by stimuli associated with food delivery [73,101]. And, indeed, in two recent series of experiments we have found clear evidence of the involvement of the BLA in incentive learning. In one series we found that lesions of the BLA rendered the instrumental performance of rats insensitive to outcome devaluation, apparently because they were no longer able to associate the sensory features of the instrumental outcome with its incentive value [16]. More recently, we have confirmed this suggestion using post-training infusions of the protein-synthesis inhibitor anisomycin [124]. It has now been well documented that both the consolidation of the stimulus-affect association that underlies fear conditioning and its reconsolidation after retrieval depends on the synthesis of new proteins in the BLA [96,109]. In a recent experiment, we first trained hungry rats to press two levers with one earning food pellets and the other a sucrose solution. After this training the rats were sated and given the opportunity for incentive learning; i.e., they were allowed to consume either the food pellets or the sucrose solution in the sated state. Immediately after this consumption phase, half of the rats were given an infusion of...
anisomycin whereas the remainder were given an infusion of vehicle. In a subsequent choice extinction test, conducted on the two levers when sated, rats in the vehicle group performed fewer responses on the lever that, in training, delivered the outcome to which they were reexposed when sated prior to the test; i.e., the standard incentive learning effect [4]. In contrast, the infusion of anisomycin completely blocked this shift in preference. To assess whether incentive learning is subject to reconsolidation involving the BLA, we gave all of the rats a second reexposure episode to either the pellets or sucrose when sated such that, if they had been first given vehicle infusion then they were now given an anisomycin infusion whereas if they were first given an anisomycin infusion they were now given a vehicle infusion. Although, again, vehicle infused rats showed reliable incentive learning, those given the anisomycin infusion performed indifferently on the two levers despite the fact that these same rats had previously shown perfectly clear evidence of incentive learning after the first episode of reexposure [124].

Previous effects of amygdala manipulations on feeding have been found to involve connections between the amygdala and the hypothalamus [101] and, indeed, it has been well reported that neuronal activity in the hypothalamus is primarily modulated by chemical signals associated with food deprivation and food ingestion, including various macronutrients [88,110,123,129]. Conversely, through its connections with visceral brain stem, midline thalamic nuclei and associated cortical areas, the hypothalamus is itself in a position to modulate motivational and nascent affective inputs into the amygdala. These inputs, when combined with the amygdala’s sensory afferents, provide the kind of associative process required to alter incentive value and points both to the associative structure and the larger neural system underlying incentive learning generally. As illustrated in Fig. 2, this structure is based on a simple feedback circuit within which the goal or reward value of a specific event is set and, indeed, can be re-set when subsequently contacted on the basis of the animals’ current internal state (see Ref. [10] for review).

4.1. Is value symbolic or somatic?

A final issue worth considering with regard to the function and representation of reward is the question of how incentive value transfers from incentive learning to a choice test in which the performance of two actions that previously delivered now differently valued outcomes is compared. Because these tests are conducted often several days after incentive learning and in extinction, there are neither any explicit internal nor external cues that the rat can use to determine choice performance. Instead, the rats must rely on their memory of specific action–outcome associations and the current relative value of the instrumental outcomes. But how is value encoded for retrieval during this test?

One theory proposes that value is retrieved through the operation of the same processes through which it was encoded. This view is perhaps best exemplified by Damasio’s [41] somatic marker hypothesis, according to which decisions based on the value of specific goals are determined by re-experiencing the emotional effects associated with contact with that goal. An alternative theory proposes that values, once determined through incentive learning, are encoded as abstract values (e.g., ‘X is good’ or ‘Y is bad’) and so are not dependent on re-experiencing the original emotional effects associated with contact with the goal and in encoding incentive value, for their retrieval (see Ref. [12] for further discussion).

We have conducted several distinct series of experiments to test these two hypotheses and, in all of these, the data suggest that, after incentive learning, incentive values are encoded abstractly and do not involve the original emotional processes that established those values during their retrieval [6–9]. For example, some time ago we examined the involvement of the gut peptide cholecystokinin (CCK) in incentive learning conducted when rats were sated; i.e., we assessed whether we could block incentive learning under satiety by blocking the (relatively) negative feedback associated with contact with food when satiated using the CCK-A antagonist devazepide [8]. Hungry rats were trained to press a lever and pull a chain with one action earning food pellets and the other a starch solution. Rats were then sated and reexposed to both the pellets and starch, one after an injection of devazepide and the other after an injection of vehicle. In a choice extinction test on the levers and chains we found that, indeed, devazepide was successful in blocking the reduction in value induced during contact with the food outcomes when sated; rats performed more of the action that, in training, delivered the outcome re-exposed under devazepide than of the other action.

We were now in a position to test whether these same emotional responses were involved in retrieval of value on test by assessing the effects of devazepide on choice performance in the test. Clearly if re-experiencing the emotional effects associated with contact with the instrumental outcomes when
satiated determines choice performance during the test, and if devazepide blocks these emotional effects, then we should anticipate that devazepide should, at the very least, produce a reduction in choice performance on test. If, however, incentive value is encoded abstractly and does not require re-experiencing the emotional state that supported its encoding, then devazepide should have no effect on test. In fact we found clear evidence for the latter prediction and against the somatic marker hypothesis in this and in several other similar studies. In contradiction of predictions from this position, incentive value requires emotional processes for encoding but appears not to require the same processes for retrieval during free choice tests.

5. Affect and arousal: Pavlovian incentives

Perhaps the most potent factor affecting addictive behavior and the one most often cited as the cause of failures to adjust to treatment is the effect that cues associated with drug delivery have on drug seeking or, in the current context, the effect that cues associated with access to specific foods have on food seeking. In fact, the idea that a stimulus associated with a positive reinforcer or reward exerts a motivational effect on behavior originates with an early study by Estes [58]. He reported that a tone paired with food elevated lever pressing by rats that had been previously reinforced with the food reward even though this response had never been trained in the presence of the tone. As this study makes clear, there are essentially three components to experiments assessing effects of this kind: A Pavlovian phase, in which the signals for reward are established; an instrumental phase, in which actions instrumental to gaining access to reward are trained; and a test phase in which the impact of the signals for reward on the performance instrumental actions is assessed. As such both the protocol for assessing the behavioral and neural determinants of this effect and the influence of reward-related cues on instrumental performance is often referred to as Pavlovian-instrumental transfer (or, simply PIT).

In Pavlovian conditioning, the unconditioned stimulus (or US) appears to be represented in terms of multiple features or components that can enter into independent associations with conditioned stimuli (CS’s). Konorski [86] was the first to articulate this idea to explain his distinction between consummatory and preparatory conditioning. As illustrated in Fig. 3, he argued that independent associations are formed between the representation of the CS and both the sensory features and the motivational properties of the US with the former mediating US-specific, consummatory responding and the latter more general, preparatory behavior. Both connections should be supposed to have an influence on the appetitive activation of the animal but in quite distinct ways; influencing affect and arousal, respectively (see Ref. [48] for review). It has been a point of some interest whether one or other of these two connections forms the basis for PIT. In fact, recent evidence suggests that transfer can be mediated by both of them.

There is no doubt that transfer can be mediated by the representation of the sensory features of the Pavlovian reinforcer. For example, there is good evidence from within-subjects designs showing that PIT can be outcome selective where transfer appears to depend on the identity of Pavlovian and instrumental reinforcers. Colwill and Motzkin [32] (see also Refs. [33,36,38]) associated one CS with food pellets and another with a sucrose solution before training the hungry rats to lever press and chain pull for these two reinforcers. When the CS’s were presented in an extinction test, the rats performed the instrumental response trained with the same reinforcer as the CS more than the action trained with the different reinforcer.

It is clear too that transfer respects the motivational relevance of the Pavlovian US. Balleine [5] exposed thirsty rats to Pavlovian pairings of a CS with either a sucrose or a sodium chloride solution before switching the motivational state to hunger by depriving the animals of food and training them to lever press for food pellets. When the CS’s were presented while the animals were lever pressing in extinction, the sucrose CS, but not the saline CS, elevated responding above the baseline level. This selective potentiation only occurred when the animals were hungry during the test, however; if they were water-deprived then the sucrose and saline CS’s produced comparable enhancements. This result shows that Pavlovian-instrumental transfer respects the relevance of the anticipated reinforcer to the motivational state of the animal on test; the sucrose solution, unlike the sodium chloride, is relevant to both hunger and thirst.

It should also be noted that in this situation the increase in lever pressing occurred in spite of the fact that the US’s predicted by the CS’s differed from the outcome earned by the rats’ instrumental actions; the sucrose CS motivated lever pressing even when this response had been trained with food pellets. In fact, when the Pavlovian US and instrumental outcome are put in conflict with respect to the test motivational state, it is the former that determines transfer. In a similar study, Dickinson and Dawson [49] established the sucrose and pellet
CS’s while the animals were hungry and also trained them to lever press for pellets at the same time as the Pavlovian conditioning. Even though lever pressing was associated with pellets, however, it was the sucrose rather than pellet CS that potentiated lever pressing when the rats were tested in extinction when thirsty. In this case, therefore, the motivational impact of the CS was more general even though its influence was sensitive to the US’s motivational relevance. The fact that both outcome specific and motivationally general forms of transfer can be observed conforms with Konorski’s [86] description of the distinct associations formed between the CS and US, and the structure of these transfer effects illustrated in Fig. 3.

At a neural level, given the arguments raised above for the involvement of the amygdala in instrumental incentive learning, the clear involvement of stimulus-affect/arousal associations in PIT would seem to implicate the amygdala in these effects too. Indeed, it has been suggested that the BLA is involved in the formation of stimulus-reward associations on the basis of evidence that lesions of the BLA attenuate conditioned place preferences for food or drugs of abuse [59,125]. More recently it has been demonstrated that lesions of the BLA produced by local injection of the excitotoxin NMDA induce deficits in second-order conditioning and in Pavlovian reinforcer devaluation [70,111,112] suggesting, more specifically, that the BLA is involved in the associative learning processes that give CS’s access to the affective value of their associated rewards. In a recent series [37] we began an assessment of the role of the BLA in PIT by comparing the effects of cell-body lesions of the BLA and CeN on outcome-specific transfer. Rats were trained to press two levers, one earning food pellets and the other sucrose solution. They were then given Pavlovian conditioning during which two auditory stimuli (i.e., a tone and a white noise) were presented, one paired with the delivery of the pellets and the other with delivery of the sucrose, before a test was conducted in which the effects on the tone and white noise stimuli on lever press performance were assessed in extinction. As has been previously reported, we found clear evidence of outcome selective transfer in sham-lesioned rats; performance was elevated over baseline on the lever that had previously earned the same outcome as that predicted by the stimulus. Performance on the other lever was unaffected. Lesions of CeN had no effect on outcome selective PIT; the results in this group were similar to those in the sham controls. In contrast, lesions of the BLA completely abolished PIT. During the test, the presentation of the stimuli failed to influence the performance on the levers; the rate of lever pressing during the stimuli did not differ from that during which the stimuli were not presented.

In direct contrast to these findings, previous experiments examining the role of the amygdala in PIT have reported that lesions of the CeN and not the BLA were effective in abolishing PIT [67,72]. One critical difference between these studies and our own, however, was that, whereas we used an outcome-selective protocol, they used a single lever design amenable to the more general motivational influence of Pavlovian cues on instrumental performance. These differences in the role of the CeN and BLA in PIT could be reconciled, therefore, if it were demonstrated that outcome-specific PIT were mediated by the BLA and general PIT by the CeN. To assess this possibility we first developed a procedure whereby we could study both the general and outcome-selective forms of PIT in a single animal. To achieve this we added to the outcome-selective protocol a third auditory CS (i.e., a clicker) paired with an appetitive US different from both of those used in instrumental training (i.e., Polycose). Although the other auditory cues were still productive of outcome-specific PIT, this third stimulus, we found, was capable of elevating both instrumental actions above baseline; i.e., of generating a general form of PIT. In a comparison of the effects of lesions of the BLA and the CeN on the outcome-specific and general forms of PIT that followed, we confirmed, using this protocol, that lesions of the BLA abolished the outcome-specific but not the general form of PIT whereas lesions of the CeN abolished general but not the outcome-specific PIT.

It is important to note that this finding, not only reconciles disparate findings in the literature, however. In addition to this service, these results also suggest that the influence of outcome-specific affective processes involving the BLA and of motivational arousal involving the CeN on instrumental performance are doubly dissociable at the level of the amygdala; i.e., that the control of performance that they exert is independent. Although, in the past, connections between the sensory and motivational features of the distributed US representation have been found to control important aspects of evaluative conditioning [30], the current results suggest that, at least with respect to the influence of reward-related cues on instrumental performance, this connection is not functional (see Fig. 3).

6. Dissociating instrumental and Pavlovian incentive processes

Evidence from PIT provides, perhaps, the strongest support for the claim that Pavlovian and instrumental conditioning share a common reward mechanism, making plausible the general claim that it is largely Pavlovian CS’s embedded in the instrumental situation that provide the motivational support for instrumental performance. Indeed, from this perspective one may go so far as to claim that it is the effect of outcome devaluation on the motivational impact of Pavlovian cues rather than on the incentive value of instrumental outcome that is responsible for the effects of this treatment on instrumental performance generally.

In contrast to predictions derived from this claim, however, a number of recent studies have found evidence that treatments that modify the effectiveness of Pavlovian cues on instrumental performance often have little or no detectable effect on instrumental outcome devaluation. In one study, for example, peripheral administration of either the D2 antagonist pimozide or the D1D2 antagonist α-flupenthixol were found to induce both a dose-dependent decrease in instrumental lever press performance for food and to attenuate the excitatory effects of a
Pavlovian CS for food on instrumental performance. Nevertheless, neither drug was found to influence the instrumental devaluation effect induced by a shift from a food deprived to a non-deprived state [52]. It appears that the changes in the incentive value of the instrumental outcome induced by devaluation treatments are mediated by a different process to that engaged by excitatory Pavlovian cues; whereas the latter appears to be dopamine dependent, the former does not.

Bertrand and colleagues have come to a very similar conclusion based on evidence that dopaminergic compounds have dissociable effects on Pavlovian-instrumental transfer and on the appetitive orofacial reactions elicited by direct intraoral infusion of foods and fluids. These latter reactions have been proposed to reflect the hedonic impact of both foods and the CS’s that predict food delivery and have been reported to be unaffected by lesions of the striatal DA input [23], the administration of pimozide [100] or the facilitation of DA transmission by either microinjection of amphetamine into the shell region of the nucleus accumbens [131] or amphetamine-induced sensitization [132]. Nevertheless, these dopaminergic manipulations were found strongly to influence the impact of Pavlovian cues on instrumental performance.

The potentiating effect of amphetamine in the accumbens shell on PIT, for example, suggests that this region of the accumbens may be involved in PIT in a manner that does not influence outcome devaluation. We have provided direct evidence for this claim in a series of experiments in which we found that selective lesions of the accumbens shell profoundly attenuated selective transfer effects produced when a CS is paired with the same reinforcer as that earned by the instrumental action but had no effect whatever on the sensitivity of rats to selective devaluation of the instrumental outcome. We have provided direct evidence for this claim in a series of experiments in which we found that selective lesions of the accumbens shell profoundly attenuated selective transfer effects produced when a CS is paired with the same reinforcer as that earned by the instrumental action but had no effect whatever on the sensitivity of rats to selective devaluation of the instrumental outcome. This study presents then a double dissociation between the effect of shell and core lesions on outcome-selective PIT and outcome-selective devaluation effects. As a consequence, the unavoidable conclusion is that these effects are mediated by anatomically and neurochemically distinct systems; that the impact of outcome devaluation cannot be explained in terms of its influence on a Pavlovian incentive process. Nor can reference to an instrumental incentive learning processes – i.e., those processes exerting strong modulatory control of performance through the encoding and retrieval of reward value – present themselves as a functional system and are likely to be integrated at a neural level. Similarly, recent work suggests that the contribution of the distinct features of Pavlovian incentive processes, as revealed in the dissociation of specific and general transfer effects described above, may contribute differentially to the initiation of goal-directed and habitual actions. Certainly, the non-selective reinforcing function of outcomes in determining the acquisition of habitual actions and the general arousing function of stimuli that predict those outcomes appear to be related. For example, the latter has been argued to act as the motivational limb of the former [104] providing, for example, the basis for what has been identified as over-responding on interval schedules [83].

Over and above these sources of functional integration, various investigators at the meeting of the Purdue Ingestive Behavior Research Center that prompted this review provided good evidence that integration at a neural systems level might be possible particularly with regard to the motivational processes that control food ingestion and pursuit. Although evidence currently suggests that incentive learning critically involves the BLA, other work suggests that the precursors of this incentive process may involve connections between primary sensory inputs and hypothalamic nuclei that underlie the animals’ ability to use food cues (e.g., sweet taste, or food viscosity) to anticipate the nutritive and caloric consequences of eating [43,90,128]. As illustrated in Fig. 2, the sensory-
nutritive connection that opens the feedback loop underlying incentive learning could well be instantiated in this manner. Also illustrated in Fig. 3, the modulation of nutrient-related excitation, both conditioned and unconditioned, by drive state has been carefully documented both behaviorally and physiologically. Thus, for example, recent work suggests that satiation is largely preabsorptive driven by early, nutrient driven negative feedback signals to the brain stem from gastric and intestinal visceral afferents that appear to involve vagal mechanoreceptors and chemoreceptors, respectively [102]. Other signals reflect the current state of energy balance as exemplified by the adipose hormone, leptin, and the pancreatic hormone, insulin. These signals enter the brain from the blood and act on receptors in the hypothalamus and elsewhere [129].

Of course, state cues do more than just modulate nutrient expectancies and can function as cues in their own right. Indeed, one interesting aspect of incentive value not emphasized above is its modulation by state cues induced by variations in satiation. Increases (or decreases) in deprivation not only increase (or decrease) the value of foods when they are contacted in that state, but also provide a signal of current value that animals can learn to use to determine food pursuit. Although we have been unable to find any evidence for the involvement of the dorsal hippocampus in instrumental conditioning, recent evidence suggests that the ventral hippocampus may play a role in the modulation of incentive value by internal state cues. For example, Davidson and colleagues have reported particularly intriguing evidence that interfering with ventral hippocampal function reduces the ability of rats to use a state of satiety to predict changes in the value of environmental stimuli [117] as well as to inhibit food intake and so regulate body weight [42]. It should be anticipated, therefore, that damage to this area would also affect the ability of state cues to control instrumental incentive value, a prediction that remains to be tested.

There are several ways in which ventral hippocampal output could exert a modulatory influence over incentive value. Perhaps the most obvious is through connections with the hypothalamus via the septal area; it has been reported that the lateral septum acts as a topographically organized relay between hippocampus and hypothalamus [106]. Another possibility lies in connections with the accumbens shell, recently implicated in food consumption based on palatability. For example, it has been reported that opiate-induced activation of the medial shell strongly increases the consumption of palatable foods, such as fats and sugars, even in satiated rats [78], an effect that, as one might expect, depends on connections with the hypothalamus. It would appear, therefore, that a ready circuit exists for the control of the precursors of the affective responses on which incentive value is based modulated by the hippocampus. Nevertheless, the exact function of this putative circuit and, indeed, its integration with that involving the basolateral amygdala has yet to be specified.

Generally, these several points of contact between the processes governing the basic motivational processes that contribute to food consumption and those contributing to food pursuit provide a number of obvious avenues for future research. Furthermore, they offer the possibility of integration not just across neural systems mediating quite diverse capacities, but also across apparently diverse functions too. It is still a matter of some dispute how the value of goals is integrated with the cognitive processes that encode action–outcome relations. At the very least, it seems likely that the solution to this problem will require an understanding of the way that food pursuit, and the costs of that pursuit, interface with the complex processes known to subserve food consumption and its regulation.

References


[56] Ding DC, Gabbott PL, Toderell D. Differences in the laminar origin of projections from the medial prefrontal cortex to the nucleus accumbens shell and core regions in the rat. Brain Res 2001;917:81–9.


Kelly RM, Strick PL. Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. Prog Brain Res 2004;143.


Kilcross S, Coutureau E. Coordination of actions and habits in the hippocampus, amygdala, dorsal striatum. Behav Neurosci 1993;107.


