Brain Reward Circuitry: Insights from Unsensed Incentives

Roy A. Wise
Behavioral Neuroscience Branch
Intramural Research Program
National Institute on Drug Abuse
National Institutes of Health
Bethesda, Maryland 20892

The natural incentives that shape behavior reach the central circuitry of motivation trans-synaptically, via the five senses, whereas the laboratory rewards of intracranial stimulation or drug injections activate reward circuitry directly, bypassing peripheral sensory pathways. The unsensed incentives of brain stimulation and intracranial drug injections thus give us tools to identify reward circuit elements within the associative portions of the CNS. Such studies have implicated the mesolimbic dopamine system and several of its afferents and efferents in motivational function. Comparisons of natural and laboratory incentives suggest hypotheses as to why some habits become compulsive and give insights into the roles of reinforcement and of prediction of reinforcement in habit formation.

The discovery by Olds and Milner (1954) that rats would learn to work for direct electrical stimulation of the brain initiated the search for the anatomical circuitry through which the normal pleasures of life establish habits that come to dominate the behavior of higher animals. It soon became apparent that lateral hypothalamic brain stimulation was not only rewarding; it was also drive inducing (Olds and Olds, 1965; Coons et al., 1965; Glickman and Schiff, 1967). Electrical stimulation of reward-related structures thus became a tool to identify anatomical substrates presumed to participate in natural motivation (Mendelson and Chorover, 1965; MacDonnell and Flynn, 1966a, 1966b; Wise, 1974) and reward (Olds and Olds, 1963; German and Bowden, 1974; Routtenberg, 1976). Inasmuch as the reward of direct brain stimulation was not detected by sight, sound, taste, smell, or touch, it provided an unsensed incentive with some degree of anatomical specificity. Intravenous (Weeks, 1962; Thompson and Schuster, 1964; Deneau et al., 1969) and intracranial (Olds et al., 1964; Phillips and LePiane, 1980; Bozarth and Wise, 1981) drug reinforcement soon offered an unsensed incentive with neurochemical specificity. These two techniques have subsequently been used extensively to characterize brain reward circuitry with respect to both its anatomy and neurochemistry. Because these laboratory incentives are not detected in the external world of the animal, they also reveal important insights into behavior motivated by natural rewards. The present paper comprises three sections. The first characterizes the elements of brain reward circuitry that have been identified by central drug injections and intracranial stimulation and that offer our best clues as to the trigger zones at which addictive drugs initiate their habit-forming actions. The second discusses possible explanations for the fact that drug reward and brain stimulation reward establish seemingly more compulsive habits than do the natural pleasures of life. The final section illustrates—again, by contrasting sensed and unsensed incentives—the fuzziness of the distinction between the “receipt” of reward and the prediction of reward.

Anatomy of Drug Reward
While there is much to learn about which dopamine neurons play roles in incentive motivation and reinforcement and there is much more to learn about the afferents to and the efferents from those dopamine neurons, a good deal is known about the brain structures and receptor subtypes at which addictive drugs trigger their habit-forming actions. This information comes in large part from studies involving intracranial drug injections that are reviewed below. The guiding assumption of such studies is that the relevant receptors for drug reward are to be found at sites where the lowest doses of microinjected drugs are rewarding. This is a fair assumption so long as care is taken to sample enough injection sites to ensure that the site of action is at the site of microinjection. The minimum controls for ensuring the validity of this assumption are “geologic” controls; unless one can demonstrate that similar injections in the regions bounding the putative site of action are not rewarding, one can never be sure that the drug is not spreading to act at a distance (Routtenberg, 1972; Wise and Hoffman, 1992). Of particular danger with hydraulic injections is that the drug spreads up the cannula shaft to a distant site of action or to the ventricular system (a pressure sink that is frequently penetrated by injection cannulae). The dangers of such spread are well illustrated by studies of the dipsogenic actions of carbachol (Routtenberg and Simpson, 1974) and angiotensin (Johnson and Epstein, 1975).

The Mesolimbic Dopamine System
A number of drugs are rewarding when injected into the nucleus accumbens where they act at mesolimbic dopamine terminals. Amphetamine, a dopamine releaser, is self-administered (Hoebel et al., 1983) and establishes conditioned place preferences (Carr and White, 1983) when injected into this region. Amphetamine injections into this region also potentiate (summate with) the rewarding effects of lateral hypothalamic brain stimulation (Colle and Wise, 1988). The dopamine uptake inhibitors nomifensine and cocaine are also self-administered into nucleus accumbens; injections into the shell are effective, whereas injections into the more dorsal and lateral core are not (Carlezon et al., 1995). Nomifensine also potentiates lateral hypothalamic brain stimulation reward by its action in this region (Carlezon and Wise, 1996b).

Cholinergic agents are rewarding when injected into the ventral tegmental area (VTA). Cytisine, a nicotinic agonist, induces conditioned place preference when in-
jected into the VTA but not when injected just dorsal to it (Museo and Wise, 1994). The cholinergic agonist carbachol causes conditioned place preference when injected into the VTA (Yeomans et al., 1985). Carbachol and the acetylcholinesterase inhibitor neostigmine are self-administered into the VTA; posterior VTA injections are most effective, and injections dorsal or lateral to the VTA are ineffective (Ikemoto and Wise, 2002). Low doses of carbachol are effective in producing conditioned place preferences when injected into the posterior but not the anterior VTA and not dorsal to the posterior VTA (Ikemoto and Wise, 2002). Carbachol activates both muscarinic and nicotinic receptors, and each type of receptor is expressed by dopaminergic neurons (Clarke and Pert, 1985; Weiner et al., 1990) and appears to contribute to carbachol’s rewarding (Ikemoto and Wise, 2002) and reward-enhancing (Yeomans and Baptista, 1997) effects in this region.

Rewarding hypothalamic brain stimulation appears to depend on trans-synaptically induced release of acetylcholine in the VTA (Yeomans et al., 1985). The axons of the mesolimbic dopamine system have high thresholds, and very few are directly activated, at traditional stimulation parameters, by rewarding hypothalamic stimulation (Yeomans, 1989; Murray and Shizgal, 1994). The bulk of the “first-stage” hypothalamic reward fibers—the reward-relevant portion of the medial forebrain bundle—that is directly depolarized by cathodal current in the lateral hypothalamic medial forebrain bundle—are thought to be caudally projecting fibers (Bielajew and Shizgal, 1986) with refractory periods in the range of 0.4–2.5 ms (Yeomans, 1979; Bielajew et al., 1982; Gratton and Wise, 1985) and conduction velocities in the range of 2–8 m/s (Bielajew and Shizgal, 1982; Murray and Shizgal, 1994). At least the major portion of these fibers is thought to synapse in the pedunculopontine or latero-dorsal tegmental nucleus, the cholinergic efferents of which relay their message back to the VTA (Yeomans et al., 1993). Hypothalamic brain stimulation reward elevates acetylcholine levels in the VTA (Rada et al., 2000), where injections of muscarinic blockers elevate reward thresholds (Yeomans et al., 1985; Kofman et al., 1990; Yeomans and Baptista, 1997). Conversely, VTA injections of acetylcholine decrease the threshold for hypothalamic brain stimulation reward (Redgrave and Horrell, 1976). The VTA cholinergic contribution to brain stimulation reward appears to involve the activation of the mesolimbic dopamine system at M5 muscarinic receptors expressed by the dopaminergic neurons of this region (Yeomans et al., 2000, 2001; Forster et al., 2002).

Mesolimbic Afferent

Opiates appear to have their strongest rewarding effects through afferents to the mesolimbic dopamine system. Mu and delta opioids are self-administered into the region of the mesolimbic dopamine cell bodies of the VTA (Bozarth and Wise, 1981; Welzl et al., 1989; Devine and Wise, 1994; David et al., 2002). The selective mu agonist DAMGO is effective at two orders of magnitude lower doses than the selective delta agonist DPDE (Devine and Wise, 1994); the two activate the mesolimbic dopamine system with the same relative potencies (Devine et al., 1993). The presumed mechanism of action of mu opioids in this region involves disinhibition of the dopamine system by inhibition of nearby GABAergic neurons that normally hold their dopaminergic neighbors under inhibitory control (Johnson and North, 1992).

GABAergic agents themselves are also self-administered into the VTA (Ikemoto et al., 1997c, 1998b; David et al., 1997). The GABA<sub>A</sub> antagonists picrotoxin and bicuculline are self-administered into the anterior VTA, while somewhat surprisingly, the GABA<sub>A</sub> agonist muscimol is self-administered into the posterior VTA; co-infusion of muscimol antagonizes the rewarding effects of anterior VTA picrotoxin injections, and, conversely, co-infusion of picrotoxin antagonizes the rewarding effects of posterior VTA muscimol injections. Self-administration of the GABA<sub>A</sub> antagonists, at least, is thought to be dopamine dependent (Ikemoto et al., 1997b; David et al., 1997). The mechanisms for these effects are not yet completely clear, because GABA<sub>A</sub> receptors are expressed not only by VTA dopamine neurons (Sugita et al., 1992) but also by the GABAergic neurons that normally inhibit the dopamine neurons (Rick and Lacey, 1994).

Microinjections of dopamine D1 antagonists in the VTA attenuate the rewarding effects of intravenous cocaine (Ranaldi and Wise, 2001), presumably by blocking the effects of dendritically released dopamine on either GABAergic (Starr, 1987; Cameron and Williams, 1993) or glutamatergic (Kalivas and Duffy, 1995) inputs to the region. Glutamatergic input to the VTA appears to offer an important link in the brain’s reward circuitry. VTA glutamate inputs arise from cortical sites including the frontal cortex. Rats will lever-press for injections of phencyclidine and other NMDA antagonists into the frontal cortex (Carlezon and Wise, 1996a), and direct electrical stimulation in this region is also rewarding (Routtenberg and Sloan, 1972; Corbett et al., 1985). Such stimulation causes glutamate release in the VTA; blockade of ionotropic glutamate receptors in VTA blocks the rewarding effects of such stimulation as well as the ability of such stimulation to elevate nucleus accumbens dopamine levels (You et al., 1998).

Mesolimbic Afferents

The mesolimbic dopamine system synapses on the shafts of the dendritic spines of GABA<sub>A</sub>-containing medium spiny neurons of nucleus accumbens (Bouyer et al., 1994). The medium spiny neurons express both D1-type and D2-type dopamine receptors (Surmeier et al., 1996), though the D1-type receptors are largely restricted to a subpopulation of medium spiny neurons expressing enkephalin and projecting to the zona reticulata of the substantia nigra, whereas the D2-type receptors are largely restricted to a subpopulation of medium spiny neurons expressing enkephalin and projecting primarily to the pallidum (Gerfen, 1992). A third subpopulation of medium spiny neurons coexpresses substance P and enkephalin and similarly coexpresses both D1-type and D2-type dopamine receptors (Surmeier et al., 1996; Aizman et al., 2000); the projection of this subpopulation is not known. Rats do not self-administer either selective D1 or D2 agonists by themselves but do self-administer a mixture of the two (Ikemoto et al., 1997a). It is tempting to suppose that they self-administer the mixture because of its actions on the subpopulation expressing both receptor subtypes, but most behavioral effects of dopaminergic agonists require cooperativity between the two receptor subtypes (Woolverton, 1986; Walters et al., 1987; Clark and...
Glutamatergic inputs from a variety of cortical sources synapse on the heads of medium spiny neurons in nucleus accumbens, and antagonists of the NMDA-type glutamate receptor are self-administered into this region. The rewarding effects of phencyclidine and other NMDA antagonists are localized to the nucleus accumbens shell; injections in the core are not effective (Carlezon and Wise, 1996a). Unlike the rewarding effects of the dopamine uptake inhibitor nomifensine and despite the fact that phencyclidine is, like nomifensine (but at higher concentrations), a dopamine uptake inhibitor (Gerhardt et al., 1987), the effects of self-administered doses of nucleus accumbens phencyclidine and other NMDA antagonists are not antagonized by dopamine receptor blockers (Carlezon and Wise, 1996a). Nucleus accumbens injections of phencyclidine and other NMDA antagonists also potentiate lateral hypothalamic brain stimulation reward (Carlezon and Wise, 1996b).

Opiates, too, are self-administered into nucleus accumbens. Morphine (Olds, 1982; David and Cazala, 2000) and the mixed mu-delta agonist methionine enkephalin (Goeders et al., 1984) are each effective, while the selective mu agonist endomorphin-1 is not (Zangen et al., 2002). While opiates in nucleus accumbens induce locomotion, their dopamine-independent actions in nucleus accumbens require an order of magnitude higher doses than do their dopamine-dependent actions in the ventral tegmental area (Kalivas et al., 1983). Interestingly, the locomotor stimulant effects of opioids in nucleus accumbens are enhanced by treatments that cause dopamine depletion or block dopamine receptor function in this region (Stinus et al., 1985, 1986).

Finally, rats self-administer the cholinergic agonist carbachol into nucleus accumbens (Ikemoto et al., 1998a). Cholinergic interneurons are sparse in this region (they comprise less than 2% of all striatal neurons), but they branch profusely and innervate the medium spiny neurons of this region (Walaas and Fonnnum, 1979; Phelps et al., 1985).

The medium spiny neurons of nucleus accumbens are the output neurons of this region; they may be viewed as a final common path for the currently identified portions of drug reward circuitry. It would appear to be depression of medium spiny neuron output—direct inhibition in the case of dopamine agonists or opiates and inhibition of excitatory input in the case of NMDA antagonists—that is common to these various drug rewards. It is not clear whether all or merely a subset of medium spiny neurons contributes to reward function.

**Reward and Compulsion**

The defining property of rewards or “reinforcers” is that they “stamp in” (Thorndike, 1898) learned associations (Pavlov, 1928) and response habits (Thorndike 1898, 1933; Skinner, 1933). The distinction between a habit and an addiction has never been a clean distinction (West, 1992; Robinson and Pritchard, 1995; Stolerman and Jarvis, 1995). With the failure of traditional dependence theory to provide a definition of addiction that covers the cases of stimulants like cocaine and nicotine (Jaffe, 1985; Wise, 1988), the current distinction between a habit and an addiction is that an addiction is a compulsive habit maintained despite harmful consequences (Jaffe, 1985; Leshner, 1999; McLellan et al., 2000). Inasmuch as it is as difficult to objectively define compulsion as it is to define addiction, the distinction between a simple habit and a compulsive habit seems more likely to be a quantitative than a qualitative distinction. Nonetheless, the attempt to characterize the mechanisms responsible for the transition from a simple habit to a compulsive habit has led to a major thrust of current work on addiction: the search for neuroadaptations that can explain the transition from habit to compulsion (e.g., Nestler, 1992, 2001; Robinson and Berridge, 1993, 2000; White and Kalivas, 1998; Berke and Hyman, 2000; Laakso et al., 2002 [this issue of Neuron]). The unsensed incentives of brain stimulation and intravenous drugs establish self-administration habits sufficiently compulsive as to qualify as addictions and offer potential insights into the neuroadaptations involved in reinforcement and addiction.

Unsensed incentives can establish very compulsive habits of seeking and ingesting. Rats will work continuously—lever-pressing at rates of several thousand responses per hour—for days to obtain direct electrical stimulation of the lateral hypothalamus and related brain regions (Olds, 1958b; Annau et al., 1974). They do so to the exclusion of other behaviors, starving themselves for the opportunity to self-stimulate if food and stimulation are concurrently available for only a limited portion of each day (Routtenberg and Lindy, 1965). Once experienced with the stimulation, rats will cross electrified grids to gain access to the lever, accepting higher shock to obtain stimulation than they are willing to accept to obtain food (even when deprived for 24 hr [Olds, 1959]).

The most obvious hypotheses as to why brain stimulation reward is so effective are (1) that they activate the reward pathway directly, bypassing synaptic barriers in sensory pathways (Wise, 1987); (2) that they activate the reward pathway powerfully, directly depolarizing a population of reward fibers within a radius of 0.25–0.5 mm (Fouriezos and Wise, 1984); and (3) that they do so with no delay of reinforcement (even a delay of 1 s between the lever-press and the delivery of reward can dramatically reduce reward effectiveness [Black et al., 1985; Fouriezos and Randall, 1997]). Rats and monkeys will work similarly compulsively for intravenous stimulants; if given unlimited access, they will self-administer intravenous injections of these drugs to the point of severe weight loss and death (Johanson et al., 1976; Bozarth and Wise, 1985). What begins as a tentative response tendency becomes a compulsive habit very quickly.

In the case of brain stimulation reward, the habit of self-administration appears to become compulsive almost immediately. The first few lever-presses that result in lateral or posterior hypothalamic stimulation are, of course, accidental. However, rats begin to respond in a focused and frenetic fashion after as few as two or three earned stimulations (Olds, 1958a). That the behavior is locked in so quickly raises serious questions as to whether any of the known neuroadaptations associated with addiction (see, for example, Nestler, 1992, 2001; Robinson and Berridge, 1993, 2000; White and Kalivas,
1998; Berke and Hyman, 2000; Hyman and Malenka, 2001)—above and beyond the neuroadaptations involved in laying down a memory trace for past drug experience—could be a necessary condition for the development of a compulsive habit from a simple habit. In the case of self-administration of intravenous stimulants or opiates, the compulsive nature of the early response habit is less dramatic. Rats appear to pursue intravenous heroin compulsively after as little as a single earned injection, but they tend to respond at rates of only two or three responses per hour because of apparent satiety (Wise et al., 1995a). Unlike brain stimulation reward, which is usually terminated abruptly 200–500 ms after its onset, drug reward decays slowly, usually by first-order kinetics. (When trains of rewarding brain stimulation reward are delivered with the slow rise and decay times of rewarding drug injections, brain stimulation, too, produces periods of satiety [Lepore and Franklin, 1992]). While animals frequently learn to respond for intravenous cocaine in the first hour or two of opportunity, the development of compulsive responding is less rapid with this drug than with heroin. Short-latency repetition of the instrumental task is infrequent in the early days of training. For example, a typical case involved an animal that responded for 1 mg/kg injections at intervals of between 14 and 55 min in the first 4 hr session. In the second session, the animal began responding more regularly, with inter-response intervals of 6–13 min after three longer latencies at the beginning of the session. By the fourth day, the animal was responding at 6 ± 2 min intervals. This degree of regularity of responding is itself an indication of compulsive behavior and is sufficient to allow confident prediction that if unlimited drug access is continued the animal will self-administer it to the point of self-induced starvation and death (Bozarth and Wise, 1985). Thus, where compulsiveness is demonstrated in tens of seconds by animals lever-pressing for brain stimulation reward and in tens of minutes in animals lever-pressing for intravenous heroin reward, it is demonstrated more on the order of days with intravenous cocaine or amphetamine self-administration. Nonetheless, one must wonder, even with cocaine and amphetamine reward, how much neuroadaptation could have taken place before the habit was stamped in to the point of compulsion.

The classic explanation for addiction is, first, that initial "recreational" use of a drug causes compensatory physiological adaptations to the drug that render the addicted brain different from the non-addicted brain and that require the user to escalate intake in order to maintain the desired effect of the drug. In this view, the neuroadaptations are thought to explain the transition from recreational to compulsive drug use. From this perspective, it should not be necessary for the animal to self-administer the drug in order to develop the adaptations that establish compulsion. Rather, compulsive self-administration should develop quickly after sufficient pre-exposure to induce the hypothesized adaptations. This prediction does not stand up very well against the facts. For example, there have been countless failed attempts to establish compulsive alcohol consumption in rodents by first establishing alcohol dependence (Lester, 1966; Falk et al., 1972; Mello, 1973; Cicero, 1980). Some success seems possible if extreme measures are taken, but it appears that alcohol must be established as a reinforcer before dependence is induced if dependence is to augment intake in such cases (Roberts et al., 2000). Opiate dependence, too, seems not a sufficient condition for establishing compulsive opiate self-administration: it has been estimated that fewer than 0.01% of patients receiving chronic opiates passively are at risk for subsequent addiction (Woods, 1990). Finally, addiction involves compulsive habits of drug self-administration. Subjects receiving drug passively do not build up the motor memories for the skilled acts of the drug-taking ritual or the procedures of drug procurement. It is the ex-user returning on the train to the place where he or she has self-administered the drug, not the ex-patient returning on the train to the hospital where he or she was treated, that experiences drug cravings and conditioned withdrawal symptoms (O'Brien et al., 1998).

Early addiction theories focused on autonomic withdrawal symptoms as evidence for the critical physiological adaptations that were the basis of compulsive drug taking; these symptoms are now known not to be necessary for compulsive drug seeking (Deneau et al., 1969; Woods and Schuster, 1971; Bozarth and Wise, 1984; Woods, 1990). Current attention focuses instead on neuroadaptations within the brain circuitry of reward itself (Koob and Bloom, 1988; Nestler, 1992; White and Kalivas, 1998; Berke and Hyman, 2000; Nestler, 2001). Neuroadaptations within the reward circuitry, though having no overt signs, could alter drug responsiveness and thus alter drug intake. This notion is supported by findings that thresholds for brain stimulation reward are higher in animals withdrawn from various drugs of abuse (Leith and Barrett, 1976; Kokkinidis et al., 1980; Kokkinidis and McCarter, 1990; Frank et al., 1988; Schulteis et al., 1995; Watkins et al., 2000). While animals that are allowed to earn moderate doses of intravenous drugs for short periods each day tend to regulate their drug intake (Pickens and Thompson, 1968; Gerber and Wise, 1989; Wise et al., 1995a, 1995b; Ranaldi et al., 1999), animals that are given access to high doses for long periods can show escalation (Ahmed and Koob, 1998) and dysregulation (Tornatzky and Miczek, 2000) of intake. One hypothesis is that prolonged high doses destabilize the mechanisms by which animals regulate limited-access drug intake (Koob and Le Moal, 2001) and that when pushed too far they, like other stress mechanisms (Sterling and Eyer, 1988; Schulkin et al., 1994), never return fully to normal. The clearest correlate of drug satiety is dopamine level in nucleus accumbens (Wise et al., 1995a, 1995b; Ranaldi and Wise, 2001).

Animals that are given repeated intermittent experimenter-administered doses of opiates or psychomotor stimulants become sensitized to these drugs, showing increased responsiveness to their locomotor stimulating (Downs and Eddy, 1932; Segal and Mandell, 1974; Bartoletti et al., 1983; Kalivas and Duffy, 1987) and rewarding (Lett, 1989; Piazza et al., 1990; Horger et al., 1990; Shippenberg and Heidbreder, 1995; Vezina et al., 1999) effects. This sensitization is long lasting and appears to involve enhanced responsiveness of the mesolimbic dopamine system or its synaptic targets (Wolf et al., 1993; Paulson and Robinson, 1995; Heidbreder et al., 1996; Nestler, 2001). Just as desensitization or tolerance of the reward circuitry has been offered as an explanation...
of escalating and compulsive drug seeking (Koob, 1996), so has its opposite, sensitization or "reverse-tolerance," been suggested to underlie the compulsive drug seeking of addiction (Robinson and Berridge, 2000).

While it is clearly the case that brain changes associated with tolerance and dependence and brain changes associated with sensitization can develop under the right circumstances, what remains to be determined is the degree to which either of these changes is necessary for motivational habits to become compulsive. The rapid onset of compulsive self-stimulation would seem to preclude any of these drug-induced long-term neuroadap-

tations as a necessary condition for compulsive drug seeking. Indeed, given the strong dosing regimens that have been used to demonstrate reliable neuroadap-
tations, one might ask the opposite question: is, perhaps, compulsive drug seeking a necessary precursor for the development of neuroadap-
tations in animals not subjected to experimenter-administered drugs? Similarly, one might ask if other motivational compulsions, such as compulsive eating, compulsive sexual activity, or compulsive gambling, are likely to affect the nervous system strongly enough to produce any of the neuroad-
tations associated with drugs of abuse. One direc-
tion that is just beginning to be explored is whether any of the known neuroadap-
tations can be established with the minimal drug treatments necessary before drug self-
administration becomes compulsive or, for that matter, before animals become behaviorally sensitized to the drugs. A second fact that is just beginning to receive attention is that increases in the tendency for compul-
sive drug-seeking behavior can grow in the absence of drug-seeking opportunity; indeed, drug seeking can be many times stronger a few weeks (Shalev et al., 2001) or a few months (Grimm et al., 2001) after the last expo-
sure to drug. A third interesting direction is the study of the effects of drug exposure on other motivated behav-

iors (Mitchell and Stewart, 1990; Harmer and Phillips, 1998; Fiorino and Phillips, 1999; Taylor and Horger, 1999). While sensed incentives may not activate the brain strongly enough to sensitize it to drugs, drug incentives may activate it strongly enough to sensitize the brain to more natural (and modest) sensed incentives.

The alternative to the view that drug-induced neuroadap-
tations make drug-seeking compulsive is that the neuroadap-
tations that differentiate the addicted from the non-addicted brain are the neuroadap-
tations associated with the learning of the drug-seeking habit. The memory of early drug experiences are stamped in by the same reinforcement process that stamps in the ordinary habits via weaker incentives. This hypothesis offers a potential explanation of the compulsiveness not only of drug self-administration, where neuroadap-
tations have been demonstrated (Robinson et al., 2001), but also of intracranial self-stimulation, where they have not. It is possible that the neuroadap-
tations of addiction are merely the neuroadap-
tations of habit formation, stamped in more strongly by drug rewards that can elevate nucleus accum-
bens dopamine levels 3- to 5-fold more than by conven-
tional rewards that tend to elevate dopamine by a factor of 1.5 or 2 (Hernandez and Hoebel, 1988; Fiorino et al., 1997; Bassareo and Di Chiara, 1999).

Reward Receipt and Reward Prediction
Comparison of the sensed rewards of food, water, and sexual interaction with the unsensed laboratory rewards of brain stimulation and intravenous and intracranial drugs illustrates the difficulty in distinguishing the actual receipt of reward from the receipt of sensory information that reward is coming (prediction of reward). How can one single out a specific event that constitutes the receipt of reward? In the case of a food reward, is the reward received when we see it, touch it, or taste it? This question is not so easily answered as common sense would have us believe.

Rewards are, in the simplest terms, the environmental incentives we tend to approach (Schneiría, 1959). More precisely, they are the environmental incentives we re-
turn to after having previously contacted them. It is the return to a reward previously experienced that is the essence of habit and addiction. This is easily understood when the reward is localized in space by one or more of the senses. Consider the case of food reward, however. Once the animal has tasted a sweet substance, it will return to it again and again. However, the return to a previously experienced reward involves the return to reward-associated landmarks as much as it involves return to the reward itself. The animal only finds the reward by approaching the environmental stimuli that point to the location of the reward. As the animal becomes experienced at foraging for food, it identifies and is guided by more and more distal stimuli that, sequentially, help the animal reach the food. Thus, the animal might first learn the smell of a given food and begin to follow the odor trail. Next, the animal might learn the sight of the plants that give off the odor in question and learn to follow the sight path until reaching the odor trail. Finally, the animal might learn the sound of the waterfall that is near the visible landmarks and follow the sound until the landmarks are visible, the sight line until the odor trail is sensed, and then the odor to the ripe and tasty portion of the plant. The sounds, sights, and smells associated with the food are clearly predictors of reward, and the efficiency of the animal increases with the identification of more and more distal predictors of reward, predictors that guide the foraging and that are important for the "error signals" that guide corrections to the foraging path.

If the sound, sight, and smell of food are predictors of reward, what is the "receipt" of reward? The widespread assumption is that the taste of a sweet substance (when hungry) or a salty substance (when sodium deficient) constitutes the receipt of reward. In the case of un-
sensed incentives—brain stimulation reward or intrave-
nous or intracranial drug reward, for example—it is a much more central event that constitutes the receipt of reward. These rewards do not activate any of the five senses (except taste, eventually, when drugs diffuse from the blood to the saliva and reach the taste buds). Yet they have the critical attribute of all rewards; they, by association, establish otherwise neutral stimuli in the environment as things to be approached. The animal trained to lever-press for brain stimulation is guided first by visual stimuli (the sight of the distal lever) and then by tactile stimuli; the stimulation itself has no locus in space. Should we then consider the taste of sweet food simply another sensory predictor of the central process?
that really constitutes the rewarding event? In part to separate the subjective and sensory experience of reward from the central events that are critical to habit formation, psychologists have come to use Pavlov’s term “reinforcement” in preference to the lay term “reward” (Wise, 1989). Is taste merely a predictor of food reinforcement?

The term “reinforcement” was first coined by Pavlov (in 1903; cited in Pavlov, 1928) to refer to the strengthening of the association between a conditioned stimulus and its unconditioned partner. Pavlov pointed out that the effectiveness of a conditioned stimulus would extinguish if not reinforced by occasional repeat pairings with its associated unconditioned stimulus. By reinforcement he meant something akin to the “stamping in” of stimulus-response associations first discussed by Thorndike (1898). Skinner (1933) and Thorndike (1933) adopted the term reinforcement to the stamping in of response habits (1937); Skinner (1937) posited two forms of reinforcement, one associated with the stamping in of stimulus-stimulus (Pavlovian) associations and one associated with the stamping in of instrumental behavior (which he called “operant” behavior). The brain mechanism of “stamping in” is, of course, the brain mechanism of learning and memory formation, and its locus in the case of habit learning and the question of whether the stamping in of stimulus-stimulus associations involves different structures than the stamping in of response-consequence associations remain matters of speculation (see, for example, White, 1996).

The clearest illustration of the reinforcement process is the stamping in of memory that occurs when a reinforcer such as brain stimulation (Huston and Mueller, 1978) or sucrose (Messier and White, 1984) is given, independent of the animal’s performance, following an unrelated learning experience. Post-trial reinforcers (and stressors) improve consolidation of memories for immediately previous events. This post-trial consolidation of learning and memory has been suggested as the essential mechanism of reinforcement, not only the reinforcement of stimulus-stimulus (Pavlovian) associations but also the reinforcement of habit-learning (Lan dauer, 1969; Pfaff, 1969).

The stamping in of food-rewarded memories appears to depend critically on postgestional consequences of food. If animals are given neutrally flavored foods and each is accompanied by an intragastric glucose load, the animals learn flavor preferences that are proportional to the associated glucose load (Le Magnen, 1959). It takes 3 or 4 days of exposure to learn these preferences, similar to the time it takes for vitamin-deficient rats to learn preferences for flavors associated with the missing vitamin (Harris et al., 1933). Thus, it would appear that the most fundamental event in the identification of sweet taste with food reward is the stamping in of the memory for ingesting sweet substances by some postgestional consequence of those substances. That is to say, the reinforcement process begins some significant time after the taste of the food; the taste of food is, like the smell and sight of food and the sounds that precede the delivery of food, a predictor of reward and not the primary reward itself.

While animals will learn habits that are rewarded by saccharin, a non-nutritive sweet substance (Sheffield and Roby, 1950), post-trial saccharin is very ineffective, relative to equally preferred concentrations of sucrose, in stamping in post-trial memories (Messier and White, 1984). Moreover, the rewarding effects of saccharin have been demonstrated only in animals that have a history of reinforcement by sweet substances with nutritive value. The first of these in a mammal’s life is mother’s milk, and rodents begin learning milk-rewarded habits from the first postnatal day (Johanson and Hall, 1979). In all probability, then, the sweet taste of saccharin is a conditioned reinforcer, as are most predictors of reward. Consider, for example, the winning of a lottery. The excitement of reward is experienced at the announcement of the winning number and the receipt of the check, not at the postgestional receipt of the food that the money eventually buys.

This is, perhaps, best illustrated with the unsensed rewards of brain stimulation and drugs. Here, the animal has only the reward-associated lever or cue light to approach; the reward itself is not sensed peripherally, either by the distance senses of sight, hearing, and smell or by the proximal senses of taste and touch. Here, the receipt of reward is concurrent with the illumination of the cue light and the click of the relay; these are synonymous with the receipt of reward if not the perceived rewarding events themselves, at least for laboratory rats.

This is an interesting issue because of the work of Schultz and colleagues, showing that the dopamine system—clearly critical for reward function—becomes increasingly responsive to reward predictors and seemingly unresponsive to the reward itself. Unit recordings in the awake animal present a complex and interesting picture of dopaminergic responsiveness to rewards. When food reward is first earned or discovered, midbrain dopaminergic neurons respond with short-latency phasic bursts of firing (Ljungberg et al., 1992; Schultz et al., 1993; Mirenowicz and Schultz, 1994). In experiments where juice near the animal’s mouth was the incentive, the neurons appeared to respond to the taste of the juice (Schultz et al., 1993; Mirenowicz and Schultz, 1994). With repeated testing, however, this phasic response becomes associated with stimuli that predict the presentation of the incentive. In experiments where a small piece of apple was the incentive, the neurons came to respond to the sight of the apple or the click of the latch to the door that hid the apple (Ljungberg et al., 1992). As the neurons began to respond to the distant signals of the rewarding event, they ceased responding to the proximal (taste) cue (Ljungberg et al., 1992; Mirenowicz and Schultz, 1994). In these studies, dopamine neurons were seen as responding to primary rewards “only when the reward occurs unpredictably, either outside of a task or during learning. By contrast, a fully predicted reward does not elicit a response in dopamine neurons, and the response is transferred to the earliest conditioned, reward-predicting stimulus” (Schultz, 1997). This interpretation has been the focus of several recent attempts to understand the role of dopamine in motivated behavior (Schultz et al., 1997; Schultz and Dickinson, 2000; O’Doherty et al., 2002; Schultz, 2002 [this issue of Neuron]).

What are the lessons from studies of unsensed incentives for the question of whether dopamine is important for the prediction rather than the receipt of reward? The
Kiyatkin et al., 2000) before a drug such as cocaine (Domesick, 1978; Mogenson et al., 1980). The rewarding contact with the approached incentive does not mark thalamic brain stimulation, and most drugs of abuse are learned incentives that arrive tens of seconds or minutes after the cue light signals the receipt of the secondary (learned) reward. That is to say, in each case it is the receipt of reward that is rewarding rather than the mere receipt of the cue or the lever press (Ljungberg et al., 1992). In the case of intravenous or intracranial drug reward, of the mesolimbic system to the nigro-striatal dopamine system—a system widely identified with motor function—sight of the lever or cue light—can be eliminated by lesions or blockade of the output nuclei because of their association with things to come; they predict the stamping in of memory by postingestional exultation, if it were objectively studied, would under-score the fact that it is the receipt of reward that arouses us most. In the human situation, it is such things as the receipt of money, the receipt of the promise—like the taste of the apple itself—is only a predictor of the receipt of the actual reward that arouses us most. In the human situation, it is such things as the receipt of money, the receipt of the promise of an assignation, or the receipt of an invitation to compete in the finals of an athletic tournament that elicits the explosive “Yes!” and that marks, as much as anything, the emotional excitement of “receiving” reward. These things are clearly rewards, but they are conditioned rewards, not primary rewards; they are rewarding only because of previous learning. These are rewards because of their association with things to come; they are rewards because they predict just as sweet taste predicts the stamping in of memory by post-ingestional glucose—something more closely linked to the survival of the individual and the species.

In the case of intravenous or intracranial drug reward, the sensed incentives—sight of the lever or cue light—are learned incentives that arise tens of seconds or perhaps minutes (Héron et al., 1994; Stathis et al., 1995; Kiyatkin et al., 2000) before a drug such as cocaine can significantly elevate extracellular dopamine. Thus, contact with the approached incentive does not mark the receipt of the primary reward. The sight of the lighted cue light signals the receipt of the secondary (learned) incentive just as does the sight of the distant apple. The click of the lever is no more or less the receipt of cocaine reward than is the click of Ljungberg et al.’s (1992) door the receipt of apple reward. That is to say, in each case it may be what the subject is “waiting for” but it—all of the things the apple itself—is only a predictor of the reinforcer, the event that stamps in memory (Lauder et al., 1969; Pfaff, 1969; Huston et al., 1974; Messier and White, 1984).

Concluding Remarks

The mesolimbic dopamine system, its cholinergic input from the brainstem, its glutamatergic input from cortical structures including the medial and occipital prefrontal cortex and amygdala, its GABAergic inputs from striatal sources, and, finally, its GABAergic efferents in nucleus accumbens (and their glutamatergic inputs from cortical structures) comprise a major portion of the endogenous circuitry through which the pleasures of the flesh come to shape the habits of animal life (Figure 1). The proximity of the mesolimbic system to the nigro-striatal dopamine system—a system widely identified with motor function—has suggested this system to be an interface between motivational and motor mechanisms (Nauta and Domesick, 1978; Mogenson et al., 1980). The rewarding effects of food, water, sexual interaction, lateral hypothalamic brain stimulation, and most drugs of abuse can be eliminated by lesions or blockade of the output neurons of nucleus accumbens or of their dopaminergic input. This system is activated trans-synaptically by the...
normal pleasures of life but can be activated directly by the laboratory rewards of intravenous drugs or electrical or chemical brain stimulation.

The activation of this system somehow serves to establish the response habits that are followed reliably by such activation, presumably by augmenting the consolidation—by “stamping in”—the still-active memory traces of the exteroceptive (reward-associated) and interoceptive (response feedback) stimuli that led to the behavior that preceded activation of the system. This stamping in is not done by the sensory events by which we identify rewards so much as by the postsynaptic—and sometimes postigestational—consequences of those sensory events. The sensations of reward are varied; most rewards or incentives can be at least partially identified with both the distal senses of vision, audition, and olfaction and the contact senses of touch and taste. However, laboratory rewards that do not activate any of the five senses can directly activate the system even more strongly than can the natural pleasures of life. Thus, the five senses themselves are responsible not for the receipt of reward but rather the prediction of the stamping-in process. Indeed, the mesolimbic dopamine system is excited as much or more by the distant sensory message that guarantees a reward or incentive is coming as by the contact sense message that a reward has arrived. This raises the possibility that the neural circuitry of learned habits is not only stamped in after rewards are received but is also primed by the stimuli that predict that rewards are coming.

Acknowledgments
I thank Yavin Shaham for comments on an earlier draft.

References


Gerfen, C.R. (1992). The neostriatal mosaic: multiple levels of compartmental organization. Trends Neurosci. 15, 133–139.


Kofman, O., McGlynn, S.M., Olmstead, M.C., and Yeomans, J.S. (1990). Differential effects of atropine, procaine and dopamine in...
the rat ventral tegmentum on lateral hypothalamic rewarding brain stimulation. Behav. Brain Res. 38, 55–68.


Rick, C.E., and Lacey, M.G. (1994). Rat substantia nigra pars reticulata neurons are tonically inhibited via GABA\(_a\), but not GABA\(_b\), receptors in vitro. Brain Res. 659, 133–137.


