Neuroeconomics and Addiction: Integrating Neuroscience, Cognition and Behavioral Economics

Author and Co-author Contact Information:

Peter T. Radu, B.A.

Samuel M. McClure, Ph.D.*

Decision Neuroscience Laboratory, Department of Psychology, Stanford University, Stanford, California, USA.

*Corresponding author. Mailing address: Jordan Hall, Building 420, 450 Serra Mall, Stanford, CA 94305, USA. E-mail: smcclure@stanford.edu. Telephone: 1-650-209-4833. Fax: 1-650-725-5699.

Glossary

* Allostasis: a principle of physiological or behavioral self-regulation whereby relevant parameters are dynamically adjusted not to achieve absolute constancy (homeostasis), but rather to optimize performance and minimize cost relative to a given set of circumstances. It is a process of achieving stability through change.

* Anhedonia hypothesis: an hypothesis implicating dopamine transmission in the subjective pleasure associated with a positive reward; downregulation of endogenous dopamine receptors after repeated drug administration causes anhedonia when drug-free, leading to ongoing use as a mood-restorative effort.
* Delay discounting: the behavioral phenomenon whereby the subjective value of a reward diminishes as a function of the delay to its receipt.

* Hyperbolic discount function: a mathematical model of delay discounting that yields dynamically inconsistent preferences between smaller rewards available sooner and larger rewards available later (i.e., after some longer delay). The primary feature of hyperbolic discounting is increasing impatience with decreasing delays such that the agent is more likely to choose a smaller, sooner reward over a larger, later alternative the closer the outcomes are to the present.

* Neuroeconomics: the study of how organisms make value-based decisions and how these decisions are expressed neurally, cognitively, and behaviorally.

**Prediction error:** a signal quantifying the discrepancy between the predicted and experienced value of a reward; reward prediction error is signaled in the brain by midbrain dopamine neurons.

* Valuation transitivity: a relationship pattern of reward-related preferences that implies a consistency in valuation between goods. Specifically, transitivity requires that if A is preferred to B and B is preferred to C then A must be preferred to C. For example, if apple is preferred to orange, and orange is preferred to banana, than apple must also be reliably preferred to banana.

**Key Words**

Dopamine, reward, neuroscience, nucleus accumbens, amygdala, ventromedial prefrontal cortex, anterior cingulate cortex, dorsolateral prefrontal cortex, Iowa Gambling Task, substance dependence

**Abstract / Synopsis**
All substances of abuse are believed to promote pathologically maladaptive behavior by hijacking common reward-related neural circuitry. Neuroeconomics, an emerging field already making substantial contributions to our understanding of this circuitry, studies how the brain computes the value of rewards and how these value estimates are expressed neurally, cognitively, and behaviorally. Neuroeconomics provides an empirically driven, computationally rigorous framework for conceptualizing addiction as an example of biased reward responding. The field employs the parameterized decision paradigms of traditional and behavioral economics, qualifying these paradigms with decision-making findings from psychology and grounding explanations in biological mechanism. This chapter presents a neuroeconomic conception of reward valuation and an overview of its neurochemical and systems-level implementation, with special attention to the computation of subjective utility. We present evidence suggesting that an array of neural and behavioral pathologies associated with addiction constitute computational flaws at four reward-relevant dimensions, which are (1) estimates of a reward’s magnitude, (2) representations of the delay to its receipt, (3) considerations of the valence of its outcome, and (4) calculations of the probability with which that outcome (or the outcome of alternative options) will occur. Implications for treatment are briefly addressed.

**List of Abbreviations**

ACC: anterior cingulate cortex  
DA: dopamine  
DLPFC: dorsolateral prefrontal cortex  
fMRI: Functional magnetic resonance imaging  
IGT: Iowa Gambling Task  
NAcc: nucleus accumbens  
VMPFC: ventromedial prefrontal cortex
Introduction

In order to ensure evolutionary fitness, organisms must be able to recognize and secure rewards—or reinforcers such as food, water, and sex—from their environment. The importance of accurate representation of reward value, for the survival of both the individual and the species, is self-evident. However, the neural and behavioral processes by which such feats are accomplished remain far more elusive. In an ever-changing and inherently uncertain environment, by what computational processes does the brain encode the value of an experienced reward and its associated cues? How is this value representation accessed and utilized to guide behavior when these predictive cues are encountered anew? Finally, how does the organism update the valuation process when predicted rewards are no longer experienced, or when they are experienced differently than remembered? Such questions are driving the rapidly growing field of neuroeconomics, which aims to elucidate the neural and psychological components of reward valuation within a unified, computationally precise framework.

Importantly for the study of drug abuse, all addictive substances have been found to act either directly or indirectly on the brain’s reward circuitry. This implies that substances of abuse motivate and initiate psychological and behavioral responses in similar fashion to natural rewards. In those individuals for whom drug use develops into an addiction, drug responding becomes exaggerated and inflexible, representing a bias in the reward valuation process. This chapter will present the hallmark features of addiction (including craving, compulsion, and sudden, cue-driven relapse) as examples of the intersection between addiction neuroscience and neuroeconomics. Viewing the cognitive, behavioral, and neural pathologies associated with addiction from a reward valuation perspective, we believe, affords researchers the insights of a variety of fields and offers a formal, unified framework from which to interpret new findings.

Introduction to Neuroeconomics
Definition of Neuroeconomics

Neuroeconomics is the study of how organisms make value-based decisions and how these decisions are expressed neurally, cognitively, and behaviorally. Though the field concerns itself with all steps of the valuation process, from stimulus perception through action selection and motor response, by far the most progress has been made in understanding the perception of rewarding stimuli and how the brain learns to assign them value. Consequently, the discussion herein focuses on these early stages of the reward valuation process and, specifically, the influences that chronic drug exposure has on their outcome.

Neuroeconomics is a relatively new field that incorporates methodologies and findings from traditional economics, neuroscience, and psychology. As such, it retains important insights from each while recognizing that a full description of value-based choice necessitates multiple simultaneous levels of analysis. The formal mathematical modeling and quantifiably testable hypotheses employed by economists provide formal decision paradigms that can be extended to a variety of circumstances surrounding reward valuation. Crucially, this computational rigor is the foundation on which the field rests; it provides the common language into which each contributing field’s influences are translated, parameterized, and subsequently utilized to model behavior. Furthermore, economic thought has provided neuroeconomics with its most critical theoretical assumption: namely, that an organism, when responding to a reward, tends to select the option that satisfies the ultimate goal of maximizing utility. This central assumption and its implications for studying reward-related decision-making are detailed in the next section.

Nevertheless, neuroeconomics refines the traditional economic approach to modeling choice in two important ways. First, incorporating research in psychology, violations of rationality are recognized and accepted as ubiquitous. Whereas normative economic models describing optimal choice assume that people are rational agents who systematically consider all options, this assumption has been repeatedly challenged by laboratory data. Instead, it is acknowledged that humans frequently rely on biases and heuristics that are promoted and shaped by prior
experience. Of course, this strategy is not always optimal. Consider, for example, a gamble in which there is a 50% chance of winning $35 and a 50% chance of losing $25. Despite its positive expected outcome, this gamble is commonly rejected; for most, the negativity associated with potentially losing money outweighs the positivity gained from winning, demonstrating the well-documented phenomenon of “loss aversion.” The reliable frequency with which individuals misjudge probabilities serves as another example. More people are afraid to fly on airplanes, for instance, than drive in cars, despite the statistical fact that accidents are far more frequent on the road. However, images of gruesome airplane crashes are readily called to mind, ultimately weighing more heavily on judgment. This “availability heuristic” is but one of many cognitive shortcuts individuals employ when rapidly making decisions. Such findings from the social sciences have provided neuroeconomists with descriptive, empirically driven accounts of human choice that are more accurate than the prescriptive assumptions of traditional economics.

Second, systems neuroscience has established that the brain functions in a parallel, distributed manner so that information is processed simultaneously by various specialized systems. For example, vision depends on integrating information from parallel “what” and “where” pathways that separately process identity and kinetic properties of visual input. A similar parallel architecture is evident in reward processing. Reward-related decision processes involve the interaction of multiple subsystems, each of which makes some necessary contribution to the value computation. Thus, although the behavioral goal can be summarized as a singular end, reaching a decision requires contributions from several specialized regions. Often these regions act in concert to produce behavior that is efficient and goal-relevant; however, they sometimes produce conflicting signals regarding stimulus valuation. These latter cases are particularly revealing for understanding mechanisms of normal and pathological decision-making.

Overall, neuroeconomics retains the formalism of economic analysis in describing value-based choice, while also recognizing that decision-making is evolutionarily driven and not
necessarily optimal. Neuroeconomic models are thus qualified and improved in descriptive power by psychological findings and grounded in biological mechanism. Integrating all three levels of analysis within a unified and computationally rigorous theoretical framework provides a powerful tool for studying reward-related behavior.

**Calculation of Utility**

Decision science rests on the crucial assumption that individuals tend to select behaviors that maximize utility. Said differently, when faced with the prospect of a reward, people seek the course of action that yields the highest perceived outcome value. This outcome value is computed as a function of the reward’s magnitude, scaled by the perceived likelihood with which it will be received and diminished by the delay to its receipt. Of course, rewards are not just passively consumed; some effort must usually be expended (or even some danger risked) to obtain them. Thus the individual must also incorporate the estimated cost of obtaining the reward, which is again qualified by probability and delay, into the equation. Overall, subjective utility can be (roughly) conceived as expected positive value minus expected cost. Options are selected according to their relative net expected benefits. By this formulation, an outcome’s *magnitude, delay, valence, and probability* constitute the four crucial components of reward value calculation.

Successfully navigating the environment requires that individuals be able to rapidly and efficiently process information with respect to their current goals. Consciously deliberating over every stimulus and its every possible outcome is far too time consuming (and cognitively costly) to be an effective behavioral strategy. Given the multiple dimensions incorporated into value computation, and the multiple neural subsystems responsible for their integration, it is perhaps unsurprising that faults in the calculation can occur during the evaluation process. Biases in such calculations, we will argue, form the basis of addictive disorders, in which outcome utility
fails to be maximized because of (i) over-estimation of the magnitude of a drug reward (and, simultaneously, under-estimation of the value of alternative, non-drug reinforcers); (ii) suboptimal delay preferences (overvaluation of immediate rewards and steep discounting of delayed rewards); (iii) relative neglect of negatively valenced outcomes associated with the use of the substance; and (iv) misestimation of the probabilities of drug and non-drug related outcome occurrences.

We begin by briefly describing the neurochemical and systems-level components of reward valuation. Of course, chronic drug exposure has well-documented deleterious effects on these neurophysiological processes, with profound behavioral consequences. Nevertheless, understanding the hypothetically optimal performance of these systems provides a standard against which to compare the pathologies characterizing drug addiction.

**Neural Mechanisms of Reward Valuation**

This section highlights the key components of the brain’s reward valuation system. We discuss current theories of reward-related dopamine (DA) transmission, as well as the cortical and sub-cortical regions to which DA projects and by which DA release is modulated. DA functioning is the subject of an enormous body of literature (see chapter 132, this volume). Our aim is not an exhaustive review, but rather to address the critical processes with respect to current theories of reward learning and reward prediction.

**Dopamine Signaling and Reward**

Over the past several decades, consensus has emerged that the neurotransmitter dopamine is crucial for reward signaling in the brain. Originating in the midbrain ventral tegmental area and substantia nigra and projecting to a variety of cortical and subcortical nuclei, DA cells respond to primary reinforcers (those that have direct consequences for evolutionary fitness, such as food, water, and sex) as well as secondary reinforcers (which have learned or
associative value, such as money). Crucially for addiction neuroscience, all addictive drugs have been found to alter DA transmission either directly or indirectly, suggesting that drugs also act as reinforcers in the brain (see chapter 132, this volume). The past two decades have seen enormous advancements in our understanding of how DA contributes to reward-based decision-making.

The initial preeminent theory regarding reward-related DA transmission was the *anhedonia hypothesis*, holding that DA release directly mediates the pleasurable effects of rewarding stimuli. Using direct brain stimulation and food reinforcement paradigms, DA antagonists were established to reliably attenuate previously reinforced behaviors such as lever pressing, ultimately leading to response extinction. This and other related lines of research led to the idea that DA release directly signals the hedonic effects of rewards and that interference with DA dampens the positive affect associated with their receipt. In the case of addiction, chronic exogenous DA release by drugs promotes homeostatic downregulation of endogenous DA efficacy, characterizing the post-intoxication “crash” and encouraging continued use to restore mood. This theory has been remarkably influential, as evidenced not only by its direct promotion of research, but also by the popular media depictions of DA (e.g., as the “pleasure chemical”) it has engendered.

Nevertheless, serious flaws with this account have been reported. Perhaps most notably, the anhedonia hypothesis fails to distinguish between reward consumption and the initiation of anticipatory physiological and motor responses. For example, inhibition of DA neurotransmission in rodents does not attenuate hedonic responses to sucrose rewards, but rather reduces the effort rodents are willing to expend to acquire them. These and other findings suggest a multi-step process in reward responding; experiencing a rewarding stimulus that signals hedonic value (“liking”) is functionally dissociable from the motivated behavioral responses initiated to obtain or experience it (“wanting”). As such, inhibiting DA selectively impairs “wanting” without seeming to adversely affect “liking.” This *incentive salience hypothesis*
of DA transmission (see chapter 142, this volume) asserts that DA imbues reward-predictive cues with incentive value, such that encountering them anew motivates the organism to initiate reward-seeking behavior. This account nicely describes aspects of the compulsivity associated with drugs of abuse, in which addicts repeatedly crave and seek out drugs, often despite a reported absence of subsequent consummatory pleasure.

A computational framework that accounts for incentive salience attributions posits DA as a *reward prediction error* signal. According to this hypothesis, phasic DA release signals the discrepancy between the predicted value of a reward and its experienced hedonic impact; this discrepancy signal retroactively updates the incentive value of the reward-predictive state. Groundbreaking studies by Wolfram Schultz and colleagues in the 1990s demonstrated that, when a rhesus monkey repeatedly experiences a reward paired with a Pavlovian conditioned stimulus, DA cell firing ceases in response to the reward itself and instead shifts to the appearance of its predictive cue. Importantly, a predicted reward that fails to be received results in a phasic decrease of DA at the time of predicted delivery, while new, unpredicted rewards amplify DA cell response. The model posits the following: encountering a state that has previously been paired with a reward results in DA release, assigning that state a predicted value and thereby mobilizing the organism to initiate appropriate reward-seeking behaviors. Upon reward obtainment, one of three updates to the predictive value signal occurs. If the reward was better than expected (*positive prediction error*), a phasic burst in DA firing imbues the predictive cue with additional salience (in proportion to the magnitude of the perceived discrepancy). Conversely, if the reward was worse than anticipated (or not delivered at all; resulting in a *negative prediction error*), phasic decreases in DA release reduce predictive salience of the cue. Finally, if experienced utility equals predicted utility, no DA is released in response to the reward itself since no error in prediction was experienced. Through repeated retroactive refinement of predicted utility signals, DA thus enables accurate reward prediction.
However, individuals rarely, if ever, make choices about rewards in isolation. Instead, people must frequently evaluate and decide between two or more qualitatively different options (for example, choosing to spend a summer afternoon watching a movie or going on a hike). The fact that humans and other animals are able to demonstrate consistent valuation transitivity (such that, if A is preferred to B and B to C, then A will be reliably preferred to C) implies the existence of a final common pathway, or common neural currency, when evaluating rewards. Studies in rodents have established that DA demonstrates all of the properties required for such a common currency. In much the same way that a $10 bill can buy both a movie ticket and transportation to a favorite hiking trail, DA responses signal the perceived subjective value of a variety of rewards on a single, common scale. This single-scale valuation mechanism has enormous implications for drugs of abuse, whose common neural effect is to increase and/or prolong synaptic DA activity. As will be explained, the result is compulsive use of the substance at the expense of other reinforcers.

**Neural Substrates of Reward Valuation**

With this computational role of DA in reward valuation in mind, we next provide a brief overview of the brain regions that have been implicated in the process, with special attention to their roles in calculating magnitude, delay, valence, and probability. As before, our goal is not to be exhaustive, but simply to re-frame these systems as interconnected components of a reward-valuation network. To do so, we start from the perspective of a two-systems model, an idea that has a well-documented history in neuroscience, psychology, and addiction alike (for example, see chapter 137, this volume). These systems—variously labeled as “automatic” and “controlled,” “habit-based” versus “planning-based,” and “impulsive” versus “reflective”—roughly correspond to mesolimbic and mesocortical projection sites, respectively. For simplicity, we employ the neutral terminology of System 1 and System 2. We describe their neural components and contributions to reward valuation in turn.
System 1: Mesolimbic DA Projection Sites

System 1 in reward processing roughly corresponds to activity of mesolimbic DA cells and the various frontal and subcortical regions to which they project, such as the amygdaloid complex and the striatum (including the caudate and the nucleus accumbens). Cognitively, these systems seem to share the common properties of supporting behavior that is rapid, automatic, effortless, and difficult to modify; collectively, various sources of evidence now indicate that they contribute to heuristic, experience-based decision-making. The ventral portion of the striatum, or nucleus accumbens (NAcc), serves as a projection site for various cortical and subcortical inputs (such as the amygdala and prefrontal cortex), integrating these various sources of information to gate goal-relevant motor responses. In neuroeconomics terms, this suggests that the NAcc assigns a measure of predicted reward magnitude in order to prioritize and initiate consummatory behavior. Accordingly, a number of studies have found that activity in NAcc scales reliably with magnitude of expected utility; human functional magnetic resonance imaging (fMRI) studies, for example, report that NAcc activity increases proportionally to the magnitude of both primary and secondary anticipated rewards. These findings imply that the striatum is a key early processing site for computing a reward’s expected utility.

Moreover, evidence suggests that the NAcc responds differentially when a reward is immediately available than when it can only be received after some delay, indicating its role in valuation over time. Evidence from behavioral economics suggests that, all things being equal, the subjective utility of a reward decreases as a function of delay to its receipt. Individuals generally prefer to receive a reward sooner rather than later, so that in some instances, smaller, immediate rewards are preferred to delayed alternatives of larger magnitude (a phenomenon known as delay discounting; choices between temporally separated outcomes are known as inter-temporal choice). In this way, temporal distance results in generally decreasing value estimates. Correspondingly, single cell recordings from the midbrain in monkeys have found
systematic decreases in DA cell firing rates as delay to a conditioned reward increases. A similar pattern of findings has been found in human subjects using fMRI. Additionally, NAcc activity correlates with impulsive inter-temporal choice: there is a direct relationship between NAcc responsiveness is and overall tendency to choose immediate over greater, delayed rewards.

Working in close concert with the striatum is the amygdala (more generally, the amygdaloid complex), a region that seems particularly crucial in marking learned reward cue valence. Though traditionally conceived as an area signaling negative valence (and particularly fear), numerous sources of evidence now suggest that the region is involved in attributing emotional salience more generally. Lesions to the amygdala, for example, impair performance on the Iowa Gambling Task (IGT), a card game requiring participants to learn the probabilities with which four decks of cards yield high or low monetary gains and losses. Two decks yield infrequent large losses but frequent small gains (good decks with positive overall expected value), while the other two incur small, frequent losses with only occasional large gains (bad decks with negative expected value). Patients with amygdala damage fail to acquire anticipatory skin conductance responses to both good and bad decks, and behaviorally they fail to avoid bad decks when playing the task. This reflects a general inability to prospectively gauge the relative valence of choice alternatives and use this affective information to guide behavior. Amygdaloid nuclei thus seem critical for learning the association between motivationally important outcomes (such as drug intoxication) and the stimuli that become valenced predictors of these motivators (such as drug paraphernalia). Glutamatergic efferents from amygdala to NAcc (and ventromedial prefrontal cortex, described below) may represent an interconnected valuation network whereby a cue is perceived, recognized as a salient predictor of a reward, and ascribed a predicted utility accordingly.

System 2: Mesocortical DA Projection Sites
In reward processing, System 2 corresponds to a network roughly corresponding to mesocortical DA pathways, consists of lateral and medial frontal and anterior cingulate cortical sites. These regions, in contrast to System 1, underlie a type of cognition that is effortful, slow, systematic, and easily disrupted (e.g., by cognitive demand or distraction). Additionally, System 2 plays a supervisory role over the efficient but sometimes error-prone responding of System 1, necessary for correcting or overriding its output. Activity in the anterior cingulate cortex (ACC), for example, has traditionally been construed as an error detection signal whereby problematic decision strategies are recognized so that cognitive control may be recruited and corrections applied. Recent research on both humans and monkeys find that lesions to the ACC (particularly the dorsal sites) do not impair simple conflict detection and response updating on Stroop or go/no-go tasks, challenging this explanation. Instead, reward responding seems to be selectively impaired, suggesting that the ACC may support the motivation to avoid mistakes (rather than simply detecting errors *per se*). Although this interpretation remains controversial, the ACC is certainly critical for monitoring behavior and enabling individuals to emit appropriate behaviors in response to rewards, particularly in challenging environments.

Densely connected to these structures is the ventromedial prefrontal cortex (VMPFC), a region (including the orbitofrontal cortex, or OFC) that also plays a supervisory role in valuation estimates. Whereas the NAcc and amygdala seem critical in determining valence and magnitude of rewards and their associated cues, evidence suggests that the VMPFC integrates these sources of information to guide response selection and track action outcome. In studies of monkeys selecting between different juice rewards, cells recorded from this region fire in anticipation of rewards at rates that positively correlate with respective subjective preference. Importantly, in the case of valuation transitivity, the cells exhibit corresponding transitivity in firing rates. Activity in these cells may thus reflect the common neural currency signals communicated by DA release.
Additionally, cells in the VMPFC seem to be responsible for updating stimulus-reward valuation contingencies when those contingencies change, as well as tracking probabilistic occurrence of both rewards and punishers. VMPFC neurons are responsible for updating preferences for previously reinforced stimuli when their reward associations are reversed, evidence of a higher-order valuation mechanism whereby reward and punishment representations are tracked and updated over the long-term. Furthermore, patients with selective lesions to VMPFC make suboptimal deck selections on the IGT by overdrawning from the bad decks. Such behavior represents ineffective incorporation of negatively valenced outcomes (and/or hyper-responsivity to positive valenced outcomes) in valuation integration processes.

In addition to the VMPFC, dorsolateral prefrontal cortex (DLPFC) represents a region where input from a variety of cortical and subcortical valuation sites converges to make goal-appropriate decisions. For example, we have reported separate valuation systems for considering intertemporal rewards in the delay-discounting task; System 1 (especially NAcc) activity responds to immediate rewards, while system 2 (notably, DLPFC) considers all intertemporal options equally. DLPFC activity is greater than NAcc activity when larger, delayed outcomes are selected. As such, DLPFC may contribute to executive cognitive processing, weighing options against their (immediate and future) probable outcomes and modulating System 1 activity to maximize utility. Such a modulatory role for DLPFC is corroborated by findings from emotion regulation paradigms, in which DLPFC downregulates amygdala activity when subjects reappraise the meaning of a negative emotional stimulus.

Collectively, System 2 represents a critical network for traditional notions of “self-control,” whereby myopic, hedonic impulses are overridden or reconstrued to optimize behavior. As will be seen, it is an important site of dysfunction among chronic drug abusers.

**Neuroeconomics and Addiction: Biased Reward Valuation**
Research has documented that, after repeated experience with a drug reward and its associated paraphernalia, responding transition from effortful and controlled (System 2) to rapid, automatic, and contextually driven behaviors (System 1). Not surprisingly, this represents a major point of concern when the individual seeks to modify or eliminate drug use. Mounting evidence suggests that a relative imbalance between the stimulus-driven output of System 1 and the controlled guidance of System 2 underlies many addictive pathologies. We next review evidence supporting a flawed reward valuation account of substance addiction, with special attention to dysregulation of DA transmission, imbalance in dual-system processing, and the computational misestimates of magnitude, delay, valence, and probability that result.

**Magnitude**

According to the diagnostic criteria of the *DSM-IV*, two of the major characteristics of substance dependence are a compulsion to take the drug and a narrowing of the behavioral repertoire (i.e., loss of interest in other behaviors and non-drug reinforcers). These symptoms likely represent aberrant utility estimates ascribed to both drugs and natural rewards. Decades of behavioral economics research on human and animal subjects has demonstrated that drug consumption is affected by the concurrent availability of other reinforcers, and that constraints on their access or changes in their cost can increase consumption of drugs. Consistent with the idea of common neural currency, these studies demonstrate that the brain tracks subjective utilities associated with various options, selecting the reward with the highest calculated magnitude. When the cost of a reward increases, its associated net subjective utility estimate typically decreases, enhancing selection of available alternatives. Importantly for a neuroeconomic assessment of drug valuation, chronic exposure to addictive drugs biases this magnitude estimation process.

David Redish at the University of Minnesota has proposed a computational model whereby this alteration may occur. In normal learning with a natural reward (such as a particular
food item), an individual eventually learns to accurately predict its subjective utility. Consequently, no DA is released upon consumption because no prediction error is experienced. By contrast, all addictive drugs increase and/or prolong DA activity in the NAcc, where value magnitudes are tracked. Cocaine and amphetamine, for example, both act directly on DA D2 reuptake pumps, resulting in direct, prolonged stimulation of reward circuitry. As such, and with each ingestion of the substance, the brain receives a positive prediction error, thereby ascribing more and more predictive salience to the cues and states predicting substance receipt. Notably, because this large DA response is pharmacologically induced, a positive prediction error occurs independently of the subjective experience afforded by drugs, which addicts often describe as devoid of pleasure. This signal, coupled with allostatic downregulation of DA receptors in response to chronic drug exposure, results in a double-edged sword of pathological magnitude estimation: (1) abnormally high salience attributions to drug cues, precipitating compulsive drug seeking, and (2) concurrent deficiencies in endogenous DA signaling, reducing the utility of non-drug (natural) rewards.

A host of evidence regarding pathological neural responses to drugs of abuse seems to support this computational frame. Positron emission tomography (PET) studies have revealed limited DA D2 receptor availability in the NAcc across a wide variety of addictions. As D2 receptors are responsive to both natural and drug rewards, and since DA release to natural rewards is a mere fraction of that to drugs, the result is blunted affective responding to positive, non-drug cues and hyper-responsivity to conditioned drug cues. Both types of responses have been demonstrated in the amygdala-NAcc network of the addicted brain. Furthermore, human imaging studies have demonstrated abnormal VMPFC (particularly OFC) activity in human addicts, again corresponding to D2 receptor availability. OFC activity is lower than that of healthy controls when the addict is in withdrawal, but exaggerated in response to drug cues or a priming dose of drug. This pathophysiological response pattern has been hypothesized to underlie both the compulsion of drug abuse (recall that VMPFC neurons have maximal
response to those rewards that are most highly subjectively valued), as well as feelings of craving, which activates VMPFC and corresponds to heightened expectations for predicted drug rewards.

Delay

Drug use, of course, is not simply a choice among present rewards of varying magnitudes and probabilities. It also involves considering long-term outcomes (both positive and negative) and incorporating these delayed representations into the value computation. Insofar as drug addiction can be construed as persistent preference for small but immediate rewards (such as drug intoxication) instead of larger but more delayed rewards (such as salvaged interpersonal relationships, improved health, and financial stability), it is unsurprising that addicts discount the future significantly more than controls. When administered a laboratory-based delay-discounting paradigm, which pits smaller, sooner sums of money against larger, later alternatives, addicts are much more likely than controls to opt for the former. This robust behavioral economic finding extends to abusers of virtually all addictive substances. Furthermore, in the domain of nicotine addiction, the extent of future discounting has been shown to correlate with addiction severity (i.e., number of cigarettes smoked daily), to predict short-term cessation outcome, and to subsequently decrease with prolonged abstinence. Additionally, future discounting increases among opiate addicts who are in withdrawal, relative to those sated with buprenorphine. These data suggest that relative inconsideration of future consequences is a hallmark feature of substance addiction and a quantifiable risk factor for prolonged use and relapse.

As outlined above, fMRI research has implicated two separate neural systems in the choice between intertemporal rewards. In general, System 1 (most notably, NAcc) responds to immediate rewards, whereas System 2 (especially lateral prefrontal areas such as DLPFC) responds equally to rewards of all delays. Furthermore, relative activity of each corresponds to
resultant behavioral choice for decisions pitting an immediate reward against a larger, delayed alternative. Evidence from a host of paradigms suggests that the relative activity of these systems is altered in addiction (see chapter 137, this volume), having detrimental implications for intertemporal reward valuation. Specifically, chronic drug abuse can lead to hyper-responsivity of the amygdala-striatal network (System 1) and/or hypo-responsivity of the lateral prefrontal network (System 2), contributing to a relative preference for immediacy. For example, cocaine addicts and alcoholics in withdrawal have demonstrated reduced blood flow and glucose metabolism in the prefrontal cortex. Furthermore, imaging studies of methamphetamine addicts performing discounting tasks suggest an imbalance in the two systems. System 2 functioning (including DLPFC) is generally compromised in these individuals, making cognitive consideration of delayed rewards especially challenging and thus predisposing choice toward immediate options. Interestingly, synaptic plasticity studies suggest that increased stimulation of DA terminals in the NAcc attenuates System 2 input to the region by inducing long-term depression of prefrontal afferents. Of great consequence for addiction research is that chronic cocaine exposure in rats has the same effect: prolonged cocaine-induced DA activation reduces PFC input to the NAcc, shifting the balance towards subcortical (System 1) inputs. Obviously, this self-perpetuating, substance induced imbalance has devastating consequences for subjective utility estimation.

**Valence**

One of the most troubling features of addiction is the continuation of drug use despite mounting negative consequences. Indeed, substance addicts seem to be unphased by these negative outcomes, failing to adequately incorporate them into subsequent drug utility estimates. Performance of substance addicts on the Iowa Gambling Task (IGT) corroborates this conclusion: relative to healthy controls, they exhibit a pattern of insensitivity to losses and hypersensitivity to gains. Chronic cannabis users, for example, tend to treat each loss as a
constant, minor negative, rather than adjusting subsequent behavior according to the magnitude of the loss. This general pattern of inattention to losses has been replicated in heroin, alcohol, and stimulant abusers alike on another task requiring integration of positive and negative feedback, the Wisconsin Card Sorting Task. In this paradigm, participants sort a series of cards that vary on a number of dimensions (e.g., color, shape, and number of items) according to rules that change at different points in the task. On the basis of feedback on accuracy of sorting (whether correct or incorrect), they must update responding to settle on new sorting rules as they change. Importantly, substance abusers reliably demonstrate perseverance in rule responding, failing to successfully incorporate negative feedback into choices.

Data suggest that this diminished response to negative outcome valence results from relative hypoactivity in the VMPFC and ACC. Indeed, some addicts perform very similarly to VMPFC lesion patients on the IGT, a population that also shows blunted learning from negative consequences. Furthermore, opiate, cocaine, marijuana, and alcohol abusers all show diminished error signaling in the ACC when performing response inhibition tasks, such as the go/no-go or Stroop paradigms. Moreover, they seem to be less aware of errors than controls: on a response inhibition task in which subjects had to press a key to indicate when they realized they had made a mistake, cocaine users identified a smaller proportion of their incorrect responses as errors. Conceptually, this lack of awareness is consistent with behavioral evidence suggesting that, with sufficient experience, drug use rituals become highly automaticized and often initiated without conscious realization. When a user is attempting to abstain, a conflict between automatic and controlled goals arises, frequently experienced cognitively as craving (for broader conceptualizations of craving, see chapters 104, 105, 106, this volume). That craving has been found to correlate with increased ACC activity across a variety of substances of abuse perhaps suggests the computation of an error signal that alerts the individual to conflicts among motivationally relevant outcomes.
Also intriguing is the extent to which incidental negative affect may influence reward valuation. That addicts frequently relapse in response to negative affect and stress is a thoroughly documented behavioral fact. During drug withdrawal, stress hormone reactivity in the amygdala increases. The resulting increase in circulating glucocorticoids have been found to potentiate DA response to nicotine in the NAcc, representing a possible neurochemical substrate by which stress and negative affect may predispose abstainers to relapse. Precise mechanisms have yet to be elucidated, but it remains an intriguing possibility that the dense anatomical connectivity between the amygdaloid complex and the NAcc constitutes a pathway whereby negatively valenced stimuli directly or indirectly alter subsequent valuation estimates of positive, drug-related stimuli.

**Probability**

Substance addicts often misestimate (or ignore) the likelihood with which outcomes will occur, which can prolong drug use and encourage risky drug-related behaviors. Needle sharing, for example, is a frequent practice among injection drug users, despite its high corresponding risk of AIDS and other diseases. Cigarette smokers, despite awareness of the health risks associated with smoking, believe that their own risk for developing chronic smoking-related diseases is below average. In the laboratory, data from the probability-discounting paradigm, in which subjects choose between small, certain rewards and larger but probabilistically uncertain alternatives, corroborate this finding. Smokers, for example, discount probabilistic rewards more steeply than non-smokers, preferring smaller but certain outcomes at relatively higher probabilities of the alternative. This finding seems to vary with amount and rate of cigarette smoking. Collectively, these anecdotal and laboratory findings suggest that substance abusers routinely underestimate the likelihood with which various probabilistic outcomes will occur.

Meta-analytic evidence suggests that risky or uncertain outcomes are tracked in the VMPFC and rostral portions of the ACC, regions that are heavily affected by chronic substance
abuse. The fact that dysfunction in these areas causes serious behavioral limitations in delay, magnitude, and valence tracking as well suggests that probability judgments may exhibit important conceptual similarities to these utility dimensions. For example, both delay and probability discounting are fit well by hyperbolic discount functions, such that preferences reliably change with both changing delays and changing probabilities. Moreover, drug users who are willing to engage in risky needle sharing discount delayed rewards more steeply than do users who indicate they are not willing to needle share, and rats with lesions to the OFC have increased preferences for both smaller, immediate rewards and smaller, certain rewards. Additionally, willingness to engage in risky drug-related behaviors that increase the probability of later negative outcomes may arise from diminished loss aversion in addicts, a phenomenon routinely demonstrated by poor performance on the IGT. The fact that performance across a variety of tasks is suboptimal underscores the interconnectedness of the reward valuation network in the brain, the effects of drugs on which are pervasive.

Interestingly, recent data suggest dorsal portions of striatum, especially the caudate, may also track outcomes on probabilistic learning tasks. This region is heavily innervated by DA neurons and may thus represent a key region of dysfunction in chronic substance use. For example, one study of cigarette smokers performing a gambling game while undergoing fMRI examined the computation of fictive error signals—that is, error signals regarding hypothetical outcomes that might have occurred but were not actually experienced. The smokers seemed to compute this error signal, localized to the bilateral caudate, but subsequently failed to integrate it into value estimates to modify behavior. Ignoring this signal led to suboptimal behavioral choices in the task. A separate fMRI study of smokers revealed that, relative to smokers not expecting to smoke, those who believed they would soon be able to smoke showed blunted response to monetary gains in the caudate. Though this remains speculative, it may be that diminished caudate activity, especially during craving, reduces fictive responding to non-drug outcomes and heightens the utility of drug-related outcomes. Smokers, for example, judge the
probability of positive smoking consequences to be higher when in a high urge state, presumably contributing to their subsequent decision to smoke. This bias in probabilistic outcome valuation perhaps underlies a noteworthy paradox in the realm of public health: cigarette smokers are frequently aware of the large health risks associated with smoking, but nonetheless perceive their own susceptibility to these risks to be less than average—and continue to smoke all the while.

**Treatment Implications**

Importantly, computational models of DA transmission do not suggest that drugs of abuse will always be selected over non-drug reinforcers. Rather, the likelihood of selecting a drug reward should depend on the magnitude of the contrast between a non-drug reward and the current value assigned to drug rewards and their cues. A computationally sound and empirically verified explanation for the limited behavioral repertoire associated with drug addiction is that natural reinforcers simply do not engage the addicted brain’s neurocircuitry enough to motivate behavior. This implies that it should be possible to motivate drug abstinence by providing non-drug alternatives of sufficient perceived utility.

To this end, a particularly effective intervention for initiating and maintaining drug abstinence has been contingency management therapy, which provides a relatively immediate and certain schedule of (non-drug) rewards for biologically verified abstinence. In direct concordance with computational models of reward valuation, the effectiveness with which contingency management reduces drug administration varies proportionally to the magnitude of the abstinence rewards (typically vouchers redeemable for prizes) and the immediacy with which they are delivered. Furthermore, research also suggests that probabilistically reinforcing abstinence (having patients draw slips from a fishbowl that provide material incentives at varying probabilities) is an effective (and more cost effective) variation on this intervention. Research into the psychological and neural mechanisms by which these types of interventions
achieve their remarkable effect sizes is limited, but some evidence suggests systematic changes in reward valuation. One study at the University of Vermont, for example, found decreased discounting for delayed rewards with longer durations of achieved abstinence during the intervention. Consequently, it is likely that contingency management promotes drug abstinence by directly ameliorating aberrant reward valuation processes, demonstrating the potential for neuroeconomics to inform treatment interventions. To date, however, no research has investigated the neural changes associated with contingency management therapy.

**Conclusions and Future Directions**

Researchers’ understanding of addiction has been greatly enhanced by conceptualizing the disorder as a pathology of reward-learning. As such, substance addiction emerges as a quintessential example of suboptimal reward-related choice. This chapter has demonstrated that chronic exposure to drugs hijacks the neural currency (DA) with which rewards are assigned value, leading to computational biases in the dimensions of magnitude, delay, valence, and probability. The consequence of this aberrant computation is a persistent behavioral imbalance in which drugs are repeatedly consumed, even at the expense of natural reinforcers. While recognizing that drug addiction is not a unitary phenomenon and ultimately depends on dysfunction among multiple neurotransmitter systems, we nonetheless contend that the processes reviewed herein highlight important similarities across nearly all drugs of abuse and represent a key underlying pathology.

Nonetheless, support for these hypotheses is incomplete, and several important issues remain. For one, little is known about how and where the brain integrates value and probability estimates to select an appropriate action and initiate motor response patterns. Do the compulsive drug behaviors of addicts—often engaged in without reported enjoyment—result from a failure to update outcome contingencies *per se*, or a failure to assign motor control to the
appropriate system (i.e., goal- or habit-based) at the time of action selection? Some theories propose that the brain arbitrates between these two systems on the basis of which has the least uncertain estimate of utility at the time of cue onset. If so, do aberrant probability estimates of drug outcomes bias the arbitration process toward the habit system, or does attenuated input from cortical areas such as VMPFC (discussed earlier) mean that the goal system cannot effectively engage motor output even if its outcome estimate is preferred?

Second, more research is needed to determine whether suboptimal utility estimates in the brain are a cause or consequence of repeated drug use, or some combination thereof. For example, repeated cocaine exposure in rats has been shown to induce neuroplasticity in ventral striatal and midbrain dopamine neurons, suggesting neural abnormalities are the result of chronic drug use. In contrast, human imaging studies have suggested that pre-existing differences in DA D2 receptor availability may underlie vulnerability to subsequent drug addiction by mediating the perceived pleasantness of stimulant intoxication. Furthermore, clinical studies suggest that high levels of striatal D2 receptor availability may protect against alcoholism in non-alcoholic subjects with a family history of alcoholism. Clearly, reward-valuation abnormalities characteristic of addiction are neither purely causal nor consequential in nature; most likely, genetically-based biological variations predispose an individual to substance experimentation and maintenance, which results in exacerbated dysregulation of associated circuitry. Though methodologically challenging, more research is required to disentangle these mechanisms, as elucidating the predispositional factors in drug abuse will greatly enhance preventative interventions.

Despite these unresolved issues, existing evidence indicates that drug abuse can be fruitfully construed in neuroeconomic terms. We anticipate that this biologically sound, computationally rigorous conceptualization of addiction will continue to inform both research and treatment interventions alike.
Cross References

104. Craving and expectancies
105. Neural correlates of craving for psychoactive drugs
106. Relation of craving and appetitive behavior
132. The mesolimbic dopamine reward system and drug addiction
137. The neural basis of decision making in addiction
142. Incentive salience and the transition to addiction

References / Further reading


**Relevant Websites**

(None)
Author Biographies

Peter T. Radu is Lab Coordinator and researcher for the Decision Neuroscience Laboratory in the Department of Psychology at Stanford University, Stanford, CA, USA. Radu earned his B.A. in psychology from Stanford University in 2009. His research has focused on the behavioral and neural coordinates of reward valuation, with particular emphasis on mechanisms underlying intertemporal choice. Specifically, he has focused on intertemporal valuation processes among addicts (esp. cigarette smokers) and laboratory interventions that reduce preference for immediate rewards. He is interested in applying the principles of behavioral and neuroeconomics to increase the cost-effectiveness of substance abuse interventions such as contingency management.

Samuel M. McClure received a B.A. from the University of Pennsylvania and his Ph.D. in neuroscience from Baylor College of Medicine. From there he completed postdoctoral training at Princeton University before moving to Stanford University as an Assistant Professor in 2007. McClure's work has combined behavioral, computational, and neuroimaging methods to investigate the neural basis of reward processing and decision-making. More recently, he has focused on the neural mechanisms of delay discounting, describing the processes by which individuals evaluate goods that are available in the future.

Tables

(None)

Figures

Figure 1. Through the lens of neuroeconomics, value-based choices such as whether to smoke cigarettes are made on the basis of four attributes: the magnitude, probability, valence, and
delay of the possible outcomes. (A) Smoking is likely if the magnitude of the perceived reward is high, likely, positively valenced, and immediate. (B) Not smoking requires acknowledging the low value of smoking’s consequences, a small probability of positive outcomes, negative effects (health-related, financial), as well as the consideration of delayed alternative rewards that are larger than the value of smoking (such as good health). Ultimately, the decision will be based on which outcome is assigned the higher total estimated value.