Computational Models of Drug Use and Addiction: A Review

Jessica A. Mollick and Hedy Kober
Yale University

In this brief review, we describe current computational models of drug-use and addiction that fall into 2 broad categories: mathematically based models that rely on computational theories, and brain-based models that link computations to brain areas or circuits. Across categories, many are models of learning and decision-making, which may be compromised in addiction. Several mathematical models take predictive coding approaches, focusing on Bayesian prediction error. Other models focus on learning processes and (traditional) prediction error. Brain-based models have incorporated prefrontal cortex, basal ganglia, and the dopamine system, based on the effects of drugs on dopamine, motivation, and executive control circuits. Several models specifically describe how behavioral control may transition from habitual to goal-directed systems, consistent with computational accounts of compromised “model-based” control. Some brain-based models have linked this to the transition of behavioral control from ventral to dorsal striatum. Overall, we propose that while computational models capture some aspects of addiction and have advanced our thinking, most have focused on the effects of drug use rather than addiction per se, most have not been tested on and/or supported by human data, and few capture multiple stages and symptoms of addiction. We conclude by suggesting a path forward for computational models of addiction.

General Scientific Summary

One of the approaches that have been used to understand addiction is computational modeling; this involves simulating key variables and, in some cases, neural circuits, that contribute to the disorder. We propose that while published models have captured some aspects of addiction and have advanced our thinking, most have not captured multiple stages and symptoms of addiction and have focused on simple forms of drug use behavior instead. We briefly review those models and suggest a path forward for computational models of addiction.

Keywords: substance use disorders, computational psychiatry, learning and decision-making, Bayesian, predictive coding

Substance use disorders (SUDs; addictions) are the most prevalent and costly psychiatric conditions, associated with lifetime prevalence of 35.3% (National Institute of Mental Health, 2007), and costs exceeding $700 billion in the United States alone (Suzuki & Kober, 2018). Clinically, SUDs are chronic, relapsing conditions, characterized by problematic drug use leading to clinically significant impairment or distress (APA, 2013). For diagnosis, patients must also report at least two of 11 symptoms, which describe risky or compulsive use, impaired control, physiological alterations, and craving. The complexity of this clinical disorder highlights the need to understand the different components that contribute to the development and maintenance of SUDs. Further, understanding the underlying neural circuitry will help us develop more effective treatments for SUDs as well as prevention strategies. Thus, it is unsurprising that many models have been proposed to explain these disorders.

As Box (1979) noted: “all models are wrong but some are useful.” This is particularly true for addiction, because it is a complex phenomenon that develops in stages, involves many symptoms, and has many underlying environmental, biological, and neuropsychological causes. Thus, models of addiction are necessarily a simplification, and yet can still describe important aspects of the disorder, and drive thinking forward. Models have ranged from purely psychological (e.g., Kavanagh, Andrade, & May, 2005; Prochaska, DiClemente, & Norcross, 1992; Ryan,
Theorize the computational models of addiction as they fall into two categories: (a) mathematical models that do not make precise mappings of model components to neural circuits, and (b) biologically based models that map components to specific brain regions or systems. These categories are related to Marr's levels of analysis (Marr, 1982), such that mathematical models are at the algorithmic level, as they describe the particular algorithm implementing the computations described. In contrast, biologically based models are at the implementation level, as they describe the specific brain hardware that makes these computations. We will explain key computational mechanisms of several models in each category and describe the evidence for each model, when available (see Table 1, for a summary).

Mathematical Models

Perhaps the earliest model is Becker's "rational theory of addiction" (e.g., Becker & Murphy, 1988). This model proposes that individuals plan to maximize their utility, by making choices to take drugs when the benefits outweigh costs, which further increases the utility of using drugs. The model also considered factors that decrease utility, including price, effort, and penalties. Bernheim and Rangel (2004) incorporated some of these elements to characterize drug use as an irrational mistake, or a divergence between choices and preferences. They further specify that drugs sensitized an individual to environmental cues that promote use, by activating a "hot" decision-making system that exaggerates positive consequences of drug use. This increases subsequent choices to use drugs, even if it is not the most rational option. This model accounts for precommitment (limiting drug exposure) and regret over drug-taking, as decisions made by the rational or "cold" system. Neither model has been applied to human data or differentiates between drug use and addiction per se.

Reinforcement learning and dopamine-based models. One category of mathematical models focuses on reinforcement learning mechanisms, and how they are changed by drug use. These models focus on the mesolimbic dopamine system, based on data suggesting that dopamine neurons encode prediction errors (PE)—the difference between expectations and outcomes (Schultz, Dayan, & Montague, 1997). Such models typically do not describe how drug-induced learning affects other brain regions, but instead focus on simulating behavioral effects of drug use. One such account suggests that drug use sensitizes the dopamine system, enhancing the attribution of incentive salience to drug cues, leading to enhanced craving and motivation to use (Berridge, 2012). Specifically, Zhang, Berridge, Tindell, Smith, and Aldridge (2009) suggested that a physiological drive state multiplies the value of a conditioned cue, without needing new learning to enhance the cue's value. The model is consistent with findings that salt deprivation increases neural signaling in the ventral pallidum (a region associated with value encoding) for a salt-predictive cue before the animal had experienced the new value of salt (Tindell, Berridge, Zhang, Peciña, & Aldridge, 2005). Recent work showing that craving has a multiplicative effect on value for food provides support for this model (Konova, Louie, & Glimcher, 2018). However, this model does not describe addiction per se, but rather drug use. Further, it is inconsistent with prediction-error (PE) based accounts (described below) that assume that new learning must occur before any changes in value.

Computational Models of Drug Use and/or Addiction

As SUDs are characterized by many brain changes in reward and motivation circuitry (Belin, Belin-Rauscent, Murray, & Everitt, 2013; Goldstein & Volkow, 2002; Koob & Volkow, 2010), most computational models for drug use and addiction focused specifically on these systems. Broadly, such computational models fall into two categories: (a) mathematical models that do not make precise mappings of model components to neural circuits, and (b) biologically based models that map components to specific brain regions or systems. These categories are related to Marr's levels of analysis (Marr, 1982), such that mathematical models are at the algorithmic level, as they describe the particular algorithm implementing the computations described. In contrast, biologically based models are at the implementation level, as they describe the specific brain hardware that makes these computations. We will explain key computational mechanisms of several models in each category and describe the evidence for each model, when available (see Table 1, for a summary).

Mathematical Models

Perhaps the earliest model is Becker's "rational theory of addiction" (e.g., Becker & Murphy, 1988). This model proposes that individuals plan to maximize their utility, by making choices to take drugs when the benefits outweigh costs, which further increases the utility of using drugs. The model also considered factors that decrease utility, including price, effort, and penalties. Bernheim and Rangel (2004) incorporated some of these elements to characterize drug use as an irrational mistake, or a divergence between choices and preferences. They further specify that drugs sensitized an individual to environmental cues that promote use, by activating a "hot" decision-making system that exaggerates positive consequences of drug use. This increases subsequent choices to use drugs, even if it is not the most rational option. This model accounts for precommitment (limiting drug exposure) and regret over drug-taking, as decisions made by the rational or "cold" system. Neither model has been applied to human data or differentiates between drug use and addiction per se.

Reinforcement learning and dopamine-based models. One category of mathematical models focuses on reinforcement learning mechanisms, and how they are changed by drug use. These models focus on the mesolimbic dopamine system, based on data suggesting that dopamine neurons encode prediction errors (PE)—the difference between expectations and outcomes (Schultz, Dayan, & Montague, 1997). Such models typically do not describe how drug-induced learning affects other brain regions, but instead focus on simulating behavioral effects of drug use. One such account suggests that drug use sensitizes the dopamine system, enhancing the attribution of incentive salience to drug cues, leading to enhanced craving and motivation to use (Berridge, 2012). Specifically, Zhang, Berridge, Tindell, Smith, and Aldridge (2009) suggested that a physiological drive state multiplies the value of a conditioned cue, without needing new learning to enhance the cue's value. The model is consistent with findings that salt deprivation increases neural signaling in the ventral pallidum (a region associated with value encoding) for a salt-predictive cue before the animal had experienced the new value of salt (Tindell, Berridge, Zhang, Peciña, & Aldridge, 2005). Recent work showing that craving has a multiplicative effect on value for food provides support for this model (Konova, Louie, & Glimcher, 2018). However, this model does not describe addiction per se, but rather drug use. Further, it is inconsistent with prediction-error (PE) based accounts (described below) that assume that new learning must occur before any changes in value.

Computational Models of Drug Use and/or Addiction

As SUDs are characterized by many brain changes in reward and motivation circuitry (Belin, Belin-Rauscent, Murray, & Everitt, 2013; Goldstein & Volkow, 2002; Koob & Volkow, 2010), most computational models for drug use and addiction focused specifically on these systems. Broadly, such computer-
| Authors                          | What does it model? | Multiple symptoms? | Stages of addiction | Symptoms modeled | Existing data simulated (if any) | Evidence (indirect) | Primary category | Secondary category | Marr’s level | Main points                                                                                                                                                                                                 |
|---------------------------------|---------------------|--------------------|---------------------|------------------|------------------|-------------------------------|-------------------|------------------|------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Redish and Johnson (2007)       | Craving             | Yes                | Yes                 | Craving          | None              | Animal data                  | Mathematical      | Decision-making  | Algorithmic       | A model of the planning system in decision-making can account for craving, based on recognition of a high-value option.                                                                                       |
| Redish, Jensen, and Johnson (2008) | Addiction           | Yes                | Yes                 | Withdrawal, relapse, incentive salience, overvaluation of drugs, impulsivity, decreased executive function | None              | Animal and human data        | Mathematical      | Decision-making  | Algorithmic       | Decision-making consists of a goal-based planning system and a habit learning process. Addiction can be explained by vulnerabilities in this decision-making system.                                          |
| Dezfouli et al. (2009)          | Drug use            | Yes                | Yes: Early to late, no withdrawal, craving | Compulsive behavior: taking drugs despite punishment, overvaluation of drugs | Yes: Use despite punishment (Vanderschuren & Everitt, 2004), blocking (Patil et al., 2007), impulsive choice | Animal data        | Mathematical      | Decision-making  | Algorithmic       | A TD-based average reinforcement learning model, proposes addiction raises the basal reward threshold, and prediction errors are calculated relative to a capped reward level.                                        |
| Simon and Daw (2012)            | Drug use            | No                 | No                  | Craving/goal directed behaviors | None              | Animal data                  | Mathematical      | Decision-making  | Algorithmic       | Drug-seeking effects may be explained by both model-based and model-free systems. Outlines architectures for potential interactions between the two systems.                                              |
| Friston (2012)                  | Drug use            | No                 | No                  | Impaired reversal learning | None              | Animal and human data        | Mathematical      | Predictive coding | Algorithmic       | Addictive behavior is a consequence of impaired perceptual learning, particularly increases in precision which impairs updating of priors with precision-based prediction errors.                                    |
| Schwartenbeck et al. (2015)     | Drug use            | No                 | No                  | Impulsivity       | None              | None                         | Mathematical      | Predictive coding | Algorithmic       | Addiction leads to decreased precision in beliefs about policies, which increases impulsivity. Drug use leads to more precise beliefs about physiological states. Models data showing that beliefs about drug effects influence craving. |
| Gu and Filbey (2017); Gu (2018) | Craving             | No                 | No: Later work extends model to abstinence | Craving, effects of withdrawal duration on craving | Yes: Effect of beliefs on craving (Gu et al., 2016), effects of withdrawal duration on craving (Li, Grimm, Hope, & Shaham, 2004) | Animal and human data | Mathematical      | Predictive coding | Algorithmic       | Economic model proposes that beliefs about drug use are rational and explained by expected utility, when benefits outweigh costs.                                                                 |
| Becker and Murphy (1988)        | Drug use            | Yes                | Yes: Early to late, recovery | Initial use, recovery | None              | Human data                   | Mathematical      | Decision-making  | Algorithmic       | Economic model proposes that drug use decisions occur from interactions of a "hot" and "cold" system. Cues may activate the "hot" decision-making system and lead to irrational drug choice.    |
| Bremhein and Rangel (2004)      | Drug use            | Yes                | Yes                 | Cue-triggered craving, relapse, recovery | None              | Animal data                  | Mathematical      | Decision-making  | Algorithmic       |                                                                                                                                                                                                          |
| Authors | What does it model? | Multiple symptoms? | Stages of addiction | Symptoms modeled | Existing data simulated (if any) | Evidence (indirect) | Primary category | Secondary category | Marr’s level | Main points |
|---------|---------------------|-------------------|---------------------|------------------|-------------------------------|-------------------|-----------------|------------------|--------------|-------------|----------------|
| Redish (2004) | Drug use | No | Yes: Early to late, no withdrawal, craving | Continued learning about drugs, overvaluation of drugs | Yes: Dopamine bursts continue to occur at drug receipt (Phillips, Stuber, Heien, Wightman, & Carelli, 2003), greater elasticity (cost-sensitivity) of drug rewards (Bickel & Marsch, 2001). | Some human data, not animal data | Mathematical | Learning | Algorithmic | Learning model, predicts that prediction errors continue to occur for drug rewards because of effects on the dopamine system. |
| Takahashi, Schoenbaum, and Niv (2008) | Drug use | No | Yes: Early to late, no withdrawal, craving | Dysfunctional learning | None | Animal data | Mathematical and brain-based | Learning: actor/ critic | Implementational | Both actor (action selection) and critic (learning/prediction error) systems contribute to decision-making. In addiction, impaired encoding of state values in the critic may capture animal data on the effects of drugs on these systems. |

**Note.** This table summarizes the published models on drug use and addiction, as reviewed in this article. Note that not all models are supported by data, and very few are supported by human data. Additionally, most of them do not model multiple stages or symptoms of addiction, but rather describe circumscribed changes associated with instances of drug use, such as the effects of drug use on learning and decision-making.
Indeed, other models have focused on the effect of drug use on dopamine, capitalizing on cocaine’s effect on the dopamine transporter (Redish, 2004). These models predict that PEs would consistently occur for drug rewards, such that cues that lead to drug rewards would continue to grow in value. In turn, the model predicts that blocking—the ability of previous learning about a cue to prevent new learning about another cue—would not occur. This model is consistent with reduced blocking in human methamphetamine users (Freeman et al., 2013), and yet inconsistent with the occurrence of blocking in drug-taking animals (Panlilio, Thordike, & Schindler, 2007) and human cigarette smokers (Freeman, Morgan, Beesley, & Curran, 2012). Additional evidence suggests that negative PEs can occur for drug rewards, as rats reduce lever pressing for a smaller-than-expected drug reward (Marks, Kearns, Christensen, Silberberg, & Weiss, 2010), in contrast to the model’s prediction that drug rewards always cause positive PEs.

**Drug taking as instrumental behavior.** Several models extend this focus on learning and PEs to understand drug-taking as instrumental behavior. One of these models uses the state-action value learning framework to model the effects of drug use on learning values for states and actions (Dezfouli et al., 2009). This model proposes that drug use raises the basal reward level, and that PEs are calculated relative to a capped reward level. This enhances PEs and the value of drug-taking actions, especially early in learning. This model simulates data whereby rats continue lever pressing for a drug even when the lever is presented with a punishment cue (Vanderschuren & Everitt, 2004). This is because drug value grows over time, allowing it to compete with the lower-valued punishment. This model makes additional predictions about extended drug use, whereby drug value plateaus, because the basal reward level slowly rises. Thus, PEs are minimized over time, reducing the value of drug-related actions (e.g., lever pressing). This reduction in drug value allows the model to simulate blocking after extended drug use, consistent with prior data (Freeman et al., 2012; Panlilio et al., 2007), and in contrast with the model proposed by Redish (2004). This model also simulates impulsivity. However, it also focuses on drug-taking, and does not incorporate many aspects of addiction. Further, the evidence regarding reward PE in addiction is mixed, such that some studies have shown differences between participants with and without SUDs (Parvaz et al., 2015; Rose et al., 2014; Tanabe et al., 2013), while others have not (Park et al., 2010; Reiter et al., 2016).

“Model-based,” “model-free,” goals, and habits. Another class of mathematical models differentiates between “model-free” and “model-based” decision-making. While neurobiological substrates have been proposed for these systems, the primary mechanisms are computational (Doll, Simon, & Daw, 2012). In “model-free” control, PEs update action values based on past experience, and these action values are saved and used to guide future action-selection. However, in “model-based” control, an internal model of the world evaluates actions based on their prospective consequences, allowing updating of action values without directly experiencing all potential outcomes (Daw, Niv, & Dayan, 2005; Doll et al., 2012). “Model-based” control has been linked to goal-directed instrumental control that is sensitive to the current outcome values, in contrast to “model-free” habitual instrumental control, whereby responding persists despite changes in outcome values (Voon, Reiter, Sebold, & Groman, 2017). It has been proposed that both systems interact to control ongoing behavior (Daw et al., 2005) and rely on value computation in ventral striatum (VS; Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Simon & Daw, 2011). Simon and Daw (2012) applied the Dyna computational framework to the interactions of “model-based” and “model-free” systems in drug use. This model suggests that the high value of actions associated with drug-reward leads to enhanced “model-based” updating of states (e.g., cue-induced reactivity) and actions (e.g., conditioned drug-seeking) that precede drug-taking, via mental simulation, which is a feature of the Dyna framework. Further, another advantage of this model is that it provides a framework for both “model-based” and “model-free” control to contribute to learning. Another advantage of this model is that it incorporates goal-directed states and craving; however, it does not discuss the different stages of addiction, or model human data.

Paralleling “model-based” and “model-free” systems, Redish, Jensen, and Johnson (2008) proposed a decision-making model of addiction focusing on the interaction of a goal-based planning system, and a habit-learning process. The planning system consists of interacting components: a recognition component that identifies the current situation, a prediction component that calculates potential action consequences, and an evaluative component that calculates the value of consequences. However, this model does not discuss the different factors leading to the deployment of habit-based or goal-directed control. In an earlier version of this model, Redish and Johnson (2007) suggested that cravings may occur when the planning system recognizes a high-value drug option. This causes a bias toward retrieving associated actions in memory, which leads to recurrent drug-seeking. Redish and Johnson (2007) further propose that the planning system component is consistent with findings that rats pause before making decisions, and that hippocampal firing during pauses reflects evaluation of different maze trajectories (Johnson & Redish, 2007). In a later iteration, Redish et al. (2008) proposed that behavioral control shifts from the planning system to the habitual system, consistent with some animal studies (e.g., Hikosaka et al., 1999; Packard & McGaugh, 1996). However, this model conflates Pavlovian and instrumental learning processes within a single planning system because it uses situational cues to calculate the consequences of future actions, and the value of those consequences (Ostlund & Balleine, 2008). Indeed, Pavlovian-instrumental transfer effects (when a conditioned stimulus invigorates an instrumental response) are not sensitive to outcome devaluation (Holland, 2004; Rescorla, 1994). Further, valuations from Pavlovian and instrumental systems are neurally dissociable (Ostlund & Balleine, 2008). In response, Redish et al. (2008) suggested adding another system to encode Pavlovian stimulus-outcome associations, which could drive actions through Pavlovian conditioning (separately from the planning system, which would then only represent instrumental action). However, it is not clear how this planning system might interact with the new Pavlovian system, which is an important aspect of addictive behaviors.

Broadly, computational models of drug use that propose “model-based” versus “model-free” components are consistent with animal models that describe addiction as a transition from goal-directed to habitual behavioral control (Everitt & Robbins, 2005; Lucantonio, Caprioli, & Schoenbaum, 2014). They are fur-
ther supported by data from a study with abstinent amphetamine users who exhibited relatively more “model-free” behavior in a two-step choice task (Voon et al., 2015), and another study suggesting impaired “model-based” control in alcohol-dependent subjects (Sebold et al., 2014). However, such models are inconsistent with evidence suggesting that drug users can and do make goal-directed or “model-based” decisions (Halbout, Liu, & Ostlund, 2016; Root et al., 2009), as can human drug users (Hogarth, 2012), and humans with SUDDs (Hogarth & Chase, 2011). These models are also inconsistent with evidence suggesting that both systems can operate in parallel (Balleine & O’Doherty, 2010). Further, evidence is lacking that the complex pattern of behaviors exhibited in human addiction can become fully habitual under any circumstances (e.g., getting up, obtaining money, going to purchase drugs, etc.). These models also do not account for craving states creating the explicit goal to get drugs, which renders future instrumental drug-seeking actions goal-directed (as they are directed toward reducing craving). Nevertheless, these approaches are useful, as they provide a way for thinking about the process that determines value estimates, and have been applied to many clinical phenomena (Voon et al., 2017).

Predictive coding models. Other models are based on Bayesian probability theory, which provides a framework for performing inference, and has also been applied to drug use and addiction. Specifically, it formalizes how prior beliefs are integrated with observed data to calculate the posterior probability of a particular hypothesis. In Bayesian inference, the prior reflects prior knowledge (before considering data), while the likelihood reflects the probability of the observed data, given the hypothesis. These two terms are multiplied to calculate the posterior probability—the probability of the hypothesis given the data and the prior, normalized by the probability of the data given all possible hypotheses. The next time data are encountered, this updated posterior probability serves as the prior (Griffiths, Kemp, & Tenenbaum, 2008; Olshausen, 2004).

Friston (2005) applied the Bayesian framework to perception and cognition and their underlying neural mechanisms in a predictive coding framework. He suggested that the brain uses an internal model of the world to generate predictions about causes of sensations; sensory samples are then tested against these predictions to update beliefs about their causes (Friston, 2005, 2009, 2010). Generally, he proposed that internal brain states and actions are selected to minimize free-energy, or the difference between prior and posterior beliefs. In turn, posterior beliefs are updated after sampling the data using Bayesian updating. In Friston’s view, the difference between the prior and posterior probabilities corresponds to Bayesian surprise, or Bayesian prediction error (BPE), which is used to update beliefs and drive learning (Friston, 2005, 2009, 2010). Friston extended this account to neural hierarchies by suggesting that higher-level brain regions predict the inputs expected in the lower level, and minimize BPEs (Friston, 2005, 2009). Thus, BPEs are an algorithmic element of this predictive coding model inasmuch as they update other levels in a hierarchy. In this framework, BPEs resulting from sensory states are weighted by their confidence, or precision (inverse variance), which is captured by the dopamine system in this formulation (Friston et al., 2012). Other Bayesian models have used signed prediction errors to update beliefs (Mathys, Daumizeau, Friston, & Stephan, 2011; see online supplemental materials for additional discussion).

Based on his own approach, Friston (2012) proposed the first Bayesian account of addiction as a natural consequence of impaired perceptual learning. Incorporating physiological states, physical states, and hidden states governing causal dynamics, Friston (2012) suggests that drug-taking—that increases dopamine (Di Chiara & Imperato, 1988)—leads to high precision. This subsequently impairs learning, precision-weighed BPEs, and subsequent updating of priors. Further, this leads to a strong expectation of reward in a previously rewarded state while ignoring sensory evidence to the contrary (e.g., inaccurate priors). Given these parameters, the model further predicts that reversal learning (updating values after contingencies change) would be greatly impaired. Consistently, this model is supported by several animal studies showing that perseverative responses are enhanced after cocaine use (Calu et al., 2007; Schoenbaum, Saddoris, Ramus, Shaham, & Setlow, 2004), and two human studies demonstrating perseverative responses in drug users (Ersche et al., 2011; Ersche, Roiser, Robbins, & Sahakian, 2008). Further, the dopamine system is known to be involved in reversal learning, as administering a D2 dopamine receptor agonist reversed performance deficits in stimulant users (Ersche et al., 2011). However, impaired reversal learning could also be because of impaired encoding of (non-Bayesian) PEs or value updating, rather than enhanced precision. Further, multiple aspects of addiction are not modeled.

Gu and Filbey (2017) recently proposed a Bayesian account focused on the effect of prior beliefs on drug craving. This model draws on the effects of drugs on the dopamine system, increasing precision (Friston et al., 2012), leading specifically to more precise beliefs about physiological states in addicted individuals. The model proposes that drug-addicted individuals form priors about drug effects, including reduction in craving. Thus, with no expectation to use, the prior would shift toward greater discomfort, leading the posterior probability to shift toward greater discomfort, or more craving, upon unexpected drug receipt. In comparison, the model predicts that when drug-addicted individuals expect and receive drugs, the posterior belief would shift toward reduction of craving and lower relative discomfort. This model simulated data showing that craving was only significantly reduced when participants both expected and received drugs (Gu et al., 2016); however, it is unclear which real-life situations are modeled in this study.

Recently, Gu (2018) extended this framework to reflect changes in craving over the abstinence period, moving toward stages of addiction and recovery. However, this model—and the notion of enhanced precision of bodily state estimates—is inconsistent with findings of impaired interoceptive insight in addiction (Ateş Çol, Sönmez, & Varadar, 2016; Bechara & Damasio, 2002; Goldstein et al., 2009; Sönmez, Kahyacık Kartal, Ateş Çol, Görgülü, & Köse Çınar, 2017; Verdejo-Garcia, Clark, & Dunn, 2012), unless such estimates are both more precise and inaccurate. Further, the idea that drug use increases precision contradicts another Bayesian predictive coding account of addiction, which proposes that individuals with addiction have decreased precision in beliefs about policies (mappings from states of the world to actions), which reflects lower confidence in reaching a goal (Schwartenbeck et al., 2015). This decreased precision makes individuals more impul-
Brain-Based Models

Several brain-based models extend the focus on learning, PEs, and action selection to the interactions between the dopamine system, prefrontal cortex (PFC), and basal ganglia—crucial circuits involved in addiction (Everitt & Robbins, 2005; Goldstein & Volkow, 2011; Koob & Volkow, 2010). These include actor-critic models (Barto, 1995; Barto, Sutton, & Anderson, 1983), whereby the actor system selects different actions with an action-selection policy, while the critic system evaluates selected actions as good or bad. The critic system does this using PEs, which in these models indicates whether outcomes were better or worse than expected (unlike BPEs).

Actor-critic models were applied to interactions between the VS and dorsal striatum (DS) in addiction, as studies suggest that DS encodes action values (the actor; Burton, Nakamura, & Roesch, 2015; O’Doherty et al., 2004), while the VS encodes values of different states that contribute to PEs (the critic). An early model by Takahashi, Schoenbaum, and Niv (2008) suggests that impaired encoding of state values (cue or contexts) by the VS critic could lead to inappropriate action selection. The model is consistent with animal findings that cocaine use disrupts encoding in rat VS, while DS encoding remains (mostly) intact (Takahashi, Roesch, Stalnaker, & Schoenbaum, 2007). Impaired VS encoding could then contribute to inappropriate PEs and, thus, impaired updating of action values in the actor system. This could allow the model to capture later stages of addiction; however, this was not explicitly simulated.

Another actor-critic model suggests that a higher learning rate for appetitive learning compared with aversive learning (that increases the value of drug cues) in the VS critic could explain compulsive addictive behavior (Piray, Keramati, Dezfouli, Lucas, & Mokri, 2010). They model data observed by Deroche-Gamonet, Belin, and Piazza (2004), where a subset of vulnerable rats continued to nose-poke for a drug reward cue even when the drug cue was paired with a shock-predictive cue. Specifically, higher appetitive learning rates caused the positive value of nose-poking to outweigh the punishment. Further, the model captures enhanced control of behavior by the actor in later stages of addiction. However, neither actor-critic model has been tested in human addiction.

A related action-selection model extended the idea that control might pass from VS to DS to consider interactions between basal ganglia circuits and the PFC (Keramati & Gutkin, 2013). This model also computes action values based on PEs. Further, the model incorporates hierarchical PFC and basal ganglia loops organized by level of abstraction, such that higher levels of the model that encode abstract knowledge about the value of different options are used for updating action values at more detailed levels.

This is based on anatomical and functional evidence suggesting that representations in rostral regions of the PFC and basal ganglia are more abstract, while those in dorsal regions are more specific (Badre, 2008; Badre & Frank, 2012; Nee et al., 2013; Verstynen, Badre, Jarbo, & Schneider, 2012). The model proposes that abstract processes control behavior in the early stages of learning, because of higher flexibility. However, abstract regions (VS) are also less precise, leading to greater uncertainty as learning progresses, allowing control to pass to the more specific or less abstract regions (DS). In support of this model, some data suggest that dopamine responses to reward cues are initially higher in VS, but shift to DS over the course of learning (Willuhn, Burgeno, Everitt, & Phillips, 2012); control of drug-seeking responses may also gradually shift from the VS to DS over the course of drug use (Corbit, Nie, & Janak, 2012; Everitt & Robbins, 2016; Vanderschuren, Di Ciano, & Everitt, 2005). The model captures early and late addiction, but does not model many addiction symptoms like withdrawal or craving.

Summary and Path Forward

Theoretical and neurobiological models describe addictions as complex, multistage diseases, characterized by disturbances in subcortical (midbrain, basal ganglia, and amygdala) and cortical (prefrontal) systems, and are often based on findings from human addiction. Conversely, computational models have often focused on limited aspects of drug-taking decisions, and have relied on animal findings, which have limited their ecological validity and utility in understanding human addiction. While each modeled aspect of drug-taking may be important, few computational models captured the complexity of SUDs, as they have not integrated both the multiple stages of addiction and the multifaceted set of symptoms occurring in SUDs. Furthermore, none of the models addressed processes associated with treatment for or recovery from addiction.

Indeed, although computational models have formalized specific functions or processes involved in drug use, they have often conceptualized addiction as continued drug-taking actions that occur when drug rewards affect learning and action selection. The latter occurs either by enhancing the value of drug cues, enhancing the value of drug-taking actions, (Berridge, 2012; Dezfouli et al., 2009; Gutkin, Dehaene, & Changeux, 2006; Keramati & Gutkin, 2013; Piray et al., 2010; Redish, 2004), or by affecting precision of estimates as in predictive coding accounts (Friston, 2012; Gu & Filbey, 2017; Schwartenbeck et al., 2015). Some models address stages of addiction by proposing that control gradually shifts toward more habitual processes (Keramati & Gutkin, 2013; Redish et al., 2008), while others allow for goal-directed processes (Redish & Johnson, 2007; Simon & Daw, 2011). Commandably, a few of the models were fit to human data, focusing on craving (Gu, 2018; Gu & Filbey, 2017) and drug valuation (Redish, 2004). Others simulated animal data with addiction-relevant features such as continued drug-taking despite punishment (Dezfouli et al., 2009; Keramati & Gutkin, 2013; Piray et al., 2010). However, humans reliably develop multiple addiction symptoms concurrently (APA, 2013), and recent reverse-translational work successfully demonstrated that some drug-taking animals also develop several addiction-like symptoms, including difficulty stopping drug intake, increased effort to take drugs, and continued use.
COMPUTATIONAL MODELS OF DRUG USE AND ADDICTION

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of individual users and is not to be disseminated broadly.

despite punishment (Belin-Rauscent, Fouyssac, Bonci, & Belin, 2016; Deroche-Gamonet et al., 2004). Thus, across species, more of these behaviors could be incorporated into—and simulated by—computational models to further link them to human addiction. Finally, while each of the models made important theoretical contributions by pointing out relevant components, the lack of integration between components makes it difficult to make clinically relevant predictions or draw clinically relevant conclusions (i.e., an intervention improving one component could potentially worsen another).

In our view, the most clinically useful models are those that integrate the best of all worlds: those that are grounded in theory, describe psychological phenomena associated with addiction (not just drug use), rely on both psychological and neurobiological data (including circuits, whenever possible), formalize the components into modules, and specify computations to allow clinically relevant predictions, preferably fit to human data. An effective model of SUDs should describe the full range of addiction symptoms, along with the progression and different stages of addiction. To date, few if any such models have been put forth for addiction. Comprehensive models of this type are promising for understanding clinical disorders because they allow us to understand how addiction arises from an interaction of underlying components, and how these contribute to symptoms, which may help tailor treatments to heterogeneous patient groups or subgroups.

One promising approach that can integrate the best features of existing models is a value-based decision-making (VBDM) framework, grounded in neurobiological circuits (Mollick & Kober, 2020). A VBDM framework in the context of addiction is similar to the “Decision Theoretic Psychiatry” framework proposed by Huys, Guitart-Masip, Dolan, and Dayan (2015). This account broadly proposes that normative, instrumental choice behavior can break down—across psychiatric disorders—along three different major fault lines: determining what problem to solve, how to solve the problem, and the effects of experience. Comparatively, a VBDM framework is more specific, in proposing component computational processes that contribute to the decision-making processes described by Huys and colleagues (2015). Furthermore, we propose a specific application of the VBDM framework to SUDs, and outline several components that are unique to these disorders. As such, this VBDM framework can be applied to form a specific theory of substance use development and maintenance (see Wellman & Gelman, 1992 for a discussion of specific theories).

Indeed, VBDM seems particularly fitting to SUDs because their clinical definition explicitly describes decisions to continue using drugs despite negative consequences (Bernheim & Rangel, 2004; Redish et al., 2008). Ideally, this framework can model stages of SUD development, whereby the initial decision to take drugs can be understood in the context of interacting brain systems (e.g., craving, control), and the computation of priors. SUDs might then develop from a series of subsequent decisions, changing the weights between different nodes in the computational model, leading to future decisions to take drugs. This idea is consistent with models that showed how drug use leads to increased valuation of drug cues and actions and decisions to use despite negative consequences (Dezfouli et al., 2009; Firay et al., 2010). Building on Redish and Johnson (2007) and Simon and Daw (2012), this VBDM framework can incorporate the maintenance of goal-directed cognition in the PFC, and enhanced incentive salience to drug cues (Berridge, 2012), which characterizes craving states, and contributes to impaired executive control in addiction (Goldstein & Volkow, 2011).

Further, a VBDM framework can also incorporate learning processes, using Pavlovian and instrumental conditioning circuits in basal ganglia and amygdala. Such a framework can also incorporate the effects of beliefs on goals to take drugs, building on Bayesian models that capture the effect of beliefs on behavior, as well as the effect of bodily states on goals and motivational states (Gu & Filbey, 2017). Building on goal-directed and habitual processes (Redish et al., 2008), as well as “model-based” and “model-free” decision-making (Daw et al., 2005; Lucantoni et al., 2014; Voon et al., 2017), we suggest that engaging executive control during decision-making depends on the incentives provided, and becomes more difficult in the face of strong cue-induced craving.

Finally, unlike existing models, a VBDM framework is well-positioned to provide some insight into the ways that people recover from addiction. For example, motivational interviewing increases treatment initiation and reduces drug use by helping patients consider their long-term goals, and the relationship between their substance use and these goals. Within the VBDM framework, it can be described as enhancing the acquisition of alternative goals to drug use, which decreases the impact of drug-seeking motivations on valuation of drug cues. Contingency management (CM) is an effective treatment (Higgins, Heil, & Lussier, 2004; Regier & Redish, 2015) that provides tangible rewards for abstinence (e.g., a voucher or money for each drug-negative urine sample). In the VBDM framework, CM can be understood as changing the incentive structure for abstinence, leading to enhanced executive control in moments of craving (Kober, Kross, Mischel, Hart, & Ochsner, 2010). In turn, this makes making the decision to abstain more likely. Further, cognitive–behavioral therapies (CBT) emphasize understanding of drug use within the context of antecedents and consequences, and provide skills training to recognize situations and states where one would be vulnerable to use, including strategies for regulating craving. Within a VBDM framework, CBT can be understood as enhancing executive control systems, allowing for more deliberative, less-impulsive decision-making. This is consistent with data showing that CBT-based strategies engage executive control systems (Kober, Mende-Siedlecki, et al., 2010) and that CBT leads to lasting changes in these systems (DeVito et al., 2012, 2017). Additionally, recent evidence suggests that mindfulness-based treatments reduce drug-use (Bowen et al., 2009, 2014); they may do so by reducing the likelihood of maintaining a craving state after seeing drug cues, which reduces substance use behavior (e.g., Elwafi, Witkiewitz, Mallik, Thornhill, & Brewer, 2013; Kober, Brewer, Height, & Sinha, 2017).

The complexity of the described VBDM framework may allow it to model human behavior and its many underlying causes and consequences. As such, this framework integrates recent neurobiological research on decision-making and addiction with features from existing computational models. In doing so, it expands their scope by considering the different stages involved in drug-taking, as well as multiple addiction symptoms, and making clinically relevant predictions. Ultimately, to be of practical use, any model of addiction would need to allow for the complexity as well as the
heterogeneity of clinical presentation in humans suffering from SUDs.

References


