DOI: 10.1111/nyas.14929

REVIEW

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

# Translational models of addiction phenotypes to advance addiction pharmacotherapy

Lara A. Ray[1,2,3,4] | Steven J. Nieto[1] | Erica N. Grodin[1]

[1]Department of Psychology, University of California at Los Angeles, Los Angeles, California, USA

[2]Department of Psychiatry and Biobehavioral Sciences, University of California at Los Angeles, Los Angeles, California, USA

[3]Shirley & Stefan Hatos Center for Neuropharmacology, University of California at Los Angeles, Los Angeles, California, USA

[4]Jane & Terry Semel Institute for Neuroscience and Human Behavior, University of California at Los Angeles, Los Angeles, California, USA

**Correspondence**
Lara A. Ray, Department of Psychology, University of California at Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095–1563, USA.
Email: lararay@psych.ucla.edu

## Abstract

Alcohol and substance use disorders are heterogeneous conditions with limited effective treatment options. While there have been prior attempts to classify addiction subtypes, they have not been translated into clinical practice. In an effort to better understand heterogeneity in psychiatric disorders, the National Institute for Mental Health Research Domain Criteria (RDoC) has challenged scientists to think beyond diagnostic symptoms and to consider the underlying features of psychopathology from a neuroscience-based framework. The field of addiction has grappled with this approach by considering several key constructs with the potential to capture RDoC domains. This critical review will focus on the efforts to apply translational models of addiction phenomenology in human clinical samples, including their relative strengths and weaknesses. Opportunities for forward and reverse translation are also discussed. Deep behavioral phenotyping using neuroscience-informed batteries shows promise for a better understanding of the clinical neuroscience of addiction and advancing precision medicine for alcohol and substance use disorders.

KEYWORDS
addiction, neuroscience, pharmacotherapy, translation, treatment

## INTRODUCTION

Substance use disorder (SUD) is a medical condition that is characterized by impaired control over substance use despite adverse consequences in psychosocial domains. In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), SUDs are diagnosed by a patient's self-reported feelings and experiences and a clinician's understanding of psychiatric phenomenology or behavioral observations. These subjective approaches can lead to diagnostic inconsistencies across patients and clinicians. Additionally, diverse phenomena are captured by an SUD diagnosis. SUD requires any combination of at least two of the 11 diagnostic criteria. This means that studies on SUD (e.g., genome-wide association studies) often involve patients with different clinical presentations of SUD. The phenomenological heterogeneity that characterizes the DSM means that a diagnostic category likely encompasses a large number of biologically distinct entities.

SUD heterogeneity has also impacted treatment development. In the case of alcohol use disorder (AUD), there are currently four Food and Drug Administration (FDA)-approved medications to treat AUD: disulfiram, acamprosate, oral naltrexone, and injectable naltrexone. These medications have only modest efficacy, and given the heterogeneity of AUD, no one medication will work for every single patient suffering from this disorder. The heterogeneity of SUD and other DSM-5 disorders may be a key reason for the withdrawal of pharmaceutical companies from drug development for mental disorders. That is, if the mechanism of action of a given compound is relevant only to about half of the patients for a potential indication (i.e., a given diagnostic category), it may be expected that this will result in many failed clinical trials and an inconsistent pattern of clinical trial outcomes.

There have been attempts over the last few decades to target interventions to specific SUD subgroups. One notable example is Project MATCH, which was one of the first large-scale alcohol studies to investigate whether behavioral treatments could be targeted using a personalized treatment approach. In that study, 10 primary matching variables (e.g., perceptions and involvement of alcohol use, cognitive impairment, gender, motivation, the severity of psychiatric

  

symptoms, etc.) were studied as moderators of three behavioral treatments—cognitive behavioral therapy, motivational enhancement therapy, and 12-step facilitation.[1] Although the three behavioral treatments showed similar effectiveness, results generally failed to show that a particular therapy could be matched to particular patients to obtain better treatment outcomes. The only effective match was between patients with low psychiatric severity and 12-step facilitation, which resulted in patients having more abstinent days than similar patients treated with cognitive-behavioral therapy. Clearly, some other variables must explain the heterogeneous nature of AUD. Nearly, 25 years after the Project MATCH study, the science underlying personalized medicine is more sophisticated. In particular, progress is being made in both neurobiology and pharmacogenetics, which may facilitate the identification of biologically based AUD subtypes and the selection of treatments to target those subtypes.[2,3]

Despite advances in neuroscience and genetic research during the past two decades, there are still no genetic or other biomarkers that reliably guide the diagnosis of SUD. This makes it difficult to link SUDs to a specific circuit or gene, and in this context, neuroscience and genetic findings have had a limited impact on DSM-5 diagnostic criteria. The fact that neuroscience and genetics have contributed little to the DSM-5 may be due to several reasons. For example, the DSM-5 is central to clinical practice, insurance reimbursement, determinations of disability and service eligibility, among others. Thus, stability in diagnostic categorization is vital for these purposes, and significant changes require substantial validation. Additionally, the DSM classification system has itself impeded progress in the areas of neuroscience relevant to AUD.

In order to address the issue of SUD heterogeneity, there have been calls for the field of psychiatry to move toward a transdiagnostic and neuroscience-based framework to foster the development of psychiatric nosology based on pathophysiology rather than clinical presentation. The Research Domain Criteria (RDoC) from the National Institute of Mental Health (NIMH) is one such initiative that is intended to advance the goal of a neuroscience-based research framework for psychiatric diseases.[4] RDoC has been influential in parsing heterogeneity and focusing on behavioral assays that are translatable and biologically informative.

While the application of RDoC to addiction phenotypes has lagged behind other mental health disorders, an Alcohol Addiction RDoC (AARDoC) was proposed as a research framework wherein specific functional domains can be prioritized.[5] The AARDoC posits the following major underlying domains of functioning in AUD: reward, stress and affect regulation, incentive salience, executive function, and social processes. While the framework was proposed in 2015 by the National Institute on Alcohol and Alcoholism (NIAAA), only a few AARDoC domains have received initial validation.[6]

As a complement to these research frameworks, the Addictions Neuroclinical Assessment (ANA) was proposed by the NIAAA as a clinical framework for the assessment of addictions.[7] The ANA captures information in three of the five AARDoC domains. In its development, the ANA leveraged deep phenotyping with factor analytic methods to construct core neurofunctional domains. The ANA posits three neurofunctional domains that can be leveraged to understand heterogeneity

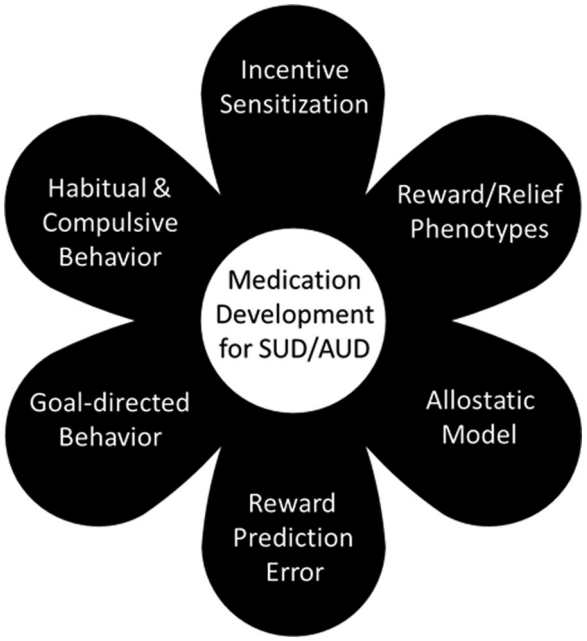


**FIGURE 1** Neuroscience-informed constructs applied to understand alcohol and substance use disorders. Neuroscience constructs reviewed include: allostasis, incentive salience, reward prediction error, reward and relief drinking phenotypes, and goal-directed, habit, and compulsive behaviors. In the area of addiction pharmacotherapy, there is a growing body of research that the reward and drinking phenotypes can inform responses to medications.

in addiction, incentive salience, negative emotionality, and executive (dys)function.[7,8] These domains have been derived across independent alcohol-focused laboratories using a combination of clinical, behavioral, and self-report measures that assess the aforementioned underlying constructs.[9–13] Our laboratory recently showed that the ANA can be applied to derive neurofunctional domain in individuals who use methamphetamine,[14] providing initial empirical support that the ANA can be applied across a range of substances.

It is possible that the ANA domains can be leveraged to advance precision medicine for AUD using pharmacological and/or behavioral treatments; however, consensus and support for a common battery are needed prior to initiating such experimental studies.

While neuroscience has yet to inform the diagnostic framework of addictive disorders, there are several neuroscience constructs that show promise in advancing pharmacotherapies for addiction (Figure 1). In an effort to advance the translational neuroscience of addictive disorders, this critical review will focus on the efforts to apply translational models of addiction phenomenology in human clinical samples, including their relative strengths and weaknesses.

## Neuroscience constructs applied to study alcohol and SUDs

### Allostasis

The allostatic model of addiction seeks to explain the intricate balance between positive and negative reinforcement in addiction.[15–17]

This model is informed by the Opponent Process Theory, developed by Solomon and Corbit (1974) to explain how two opposing processes may occur simultaneously and jointly effect motivation.[18] In basic terms, it contends that over time, addiction becomes less about positive reinforcement (the activational process, or the *a*-process) and more about negative reinforcement (the counteradaptive opponent process, or the *b*-process).[19] This theory seeks to capture the dynamic nature of addiction neurobiology as the brain is continuously adapting to large amounts of alcohol/drug use over extended periods of time, thereby causing a shift in the allostatic set point. Addiction allostasis is defined as the process of maintaining reward function stability through changes in brain reward mechanisms.[17] During the reinforcing effects of alcohol intoxication (*a*-process), there is an increase in GABAergic activity, opioid peptides, and dopamine output in the ventral striatum, which represent the neural substrates of alcohol reward. Conversely, during the counteradaptive opponent process marked by negative affect and withdrawal, there is an increase in corticotropin-releasing factor (CRF) activity as well as a decrease in neuropeptide Y (NPY), both of which are key neuromodulators of stress reactivity.[17,20,21] These processes provide the neural basis of reward and negative reinforcement associated with alcohol intoxication and withdrawal, respectively. Over time, the shift in the balance from positive to negative reinforcement is thought to explain what patients describe in their experiences with alcohol or drugs. In other words, patients often describe using drugs/alcohol to feel "normal," which is consistent with the neuroadaptation in the brain reward circuitry leading to a chronic deviation of the brain reward set point proposed by the allostatic model.[15] In this context, it may be useful for clinicians and patients alike, to recognize that from a biological standpoint, chronic and heavy alcohol use causes individuals to drink primarily to alleviate withdrawal and its associated unpleasant affective and physical symptoms.

The neurobiological underpinnings of alcohol withdrawal include changes in the neurochemical systems within the extended amygdala, including decreases in neurotransmitter functions subserving the acute reinforcing effects of alcohol (e.g., opioidergic, dopaminergic, and GABAergic).[16] An increase in alcohol self-administration can be reliably induced in animal models using a withdrawal state, and such models demonstrated that dopaminergic function is compromised during acute withdrawal.[22] Animal models have also emphasized the role of dysregulation in the brain stress system, including CRF-mediated processes, to changes in reward function leading to negative reinforcement.[17,20,21] Clinical research has sought to characterize reward and relief processes in clinical samples. In a host of laboratory studies in which individuals receive a standard dose of alcohol, our group found that the positive reinforcing effects of alcohol (i.e., stimulation and positive mood) were associated with alcohol craving in heavy drinkers but not in individuals with alcohol dependence.[23] These findings were extended in an independent sample, and it was demonstrated that the reinforcing effects of alcohol were salient determinants of subjective craving in lagged models of subjective response to alcohol and subsequent subjective craving across a range of light drinkers, heavy drinkers, and drinkers with AUD.[24] In a study combining alcohol administration with alcohol self-administration using a progressive ratio model, we did not observe a relationship between AUD severity and subjective response. AUD severity was associated with greater baseline negative mood, sedation, and craving, but it did not moderate the relationship between subjective response and subsequent self-administration. This finding did not support our prediction based on the allostatic model that higher severity individuals would show self-administration of alcohol-driven by negative reinforcement, while lower severity individuals would self-administer for positive reinforcement.[25] A longitudinal study, including alcohol administration in the laboratory, found support for the early stage phase of the allostatic model. Heavy drinkers exhibiting heightened reward sensitivity and stimulation in response to alcohol were more likely to progress to AUD.[26] In brief, studies of the subjective rewarding effects of alcohol and negative affect relief have generally provided support for predictions from the allostatic model but only for the early stage of the model. The "dark side" of addiction proposed by the model has not been reliably supported, in large part because experimental studies rely heavily on healthier and younger individuals with more mild AUD presentations.[27,28] The disconnect between AUD severity in experimental psychopathology research samples and treatment research samples has been identified as a crucial opportunity for more efficient translation, whereby clinical samples ought to express the pathology severity necessary to fully access the target constructs described in neurobiological models.[29]

Translation of the allostatic model to clinical samples with AUD has received increased attention. In the treatment domain, it has been argued that the clinical response to naltrexone, an opioid antagonist, may be stronger among individuals who report more positive reinforcement for alcohol (discussed in further detail below under reward/relief drinking).[30,31] This line of inquiry overlapped with efforts to identify genetic markers of susceptibility for alcohol reward and naltrexone responsivity in the lab[32,33] and in the clinic.[34–36] While this line of research did not produce the reliable effects required to predict naltrexone response based on the reward-drinking genotype, recent studies using self-report measures of drinking motives have shown that the effect size for treatment with naltrexone is significantly higher among individuals who report drinking for positive reinforcement, compared to individuals who drink to alleviate negative feelings or for normalization.[37,38] Collectively, these findings suggest a potential convergence between neurobiologically informed phenotypes and clinical phenotypes that can effectively improve treatment for AUD in a meaningful fashion. In medication development efforts, it has become increasingly accepted that some medications may work to attenuate the positive reinforcing effects of alcohol (e.g., naltrexone for blocking positive reinforcement), while others may be effective in targeting negative reinforcement mechanisms (e.g., gabapentin, an anticonvulsant GABA modulator, for alleviating protracted withdrawal symptoms and/or mifepristone, a glucocorticoid receptor antagonist, works on the stress system by regulating the amygdala). Alpha-1 blockers like prazosin and doxazosin may help to normalize these stress system changes seen in addictive disorders.[39] This framework continues to garner support and inform precision medicine efforts.

## Incentive sensitization

Another prominent theory of addiction consists of the incentive sensitization model. The basic tenets of this theory argue that drugs of abuse share the ability to alter brain organization (i.e., produce neuroadaptation) in the brain reward systems rendering the system sensitized to drugs and associated stimuli.[41] A key contribution of this theory is the dissociation between two aspects of incentive sensitization, namely, *liking* and *wanting*. Specifically, it has been nicely demonstrated that sensitization operates primarily at the subcomponent of reward termed incentive salience, which is marked by drug wanting. While the neural basis of liking is primarily subserved by the endogenous opioid system, the process of wanting has been associated with dopaminergic activity in the brain's reward circuitry.[40] Sensitization is not simply an inevitable pharmacological consequence of repeated drug use but instead is modulated by environmental factors associated with alcohol and drug intake.[41]

The treatment implications of incentive salience are multiple. From a neurobiological standpoint, teaching patients to cope with triggers is akin to training one's brain to unlearn associations or at a behavioral level, to inhibit a prepotent (learned) response, such as alcohol use in the presence of a drinking buddy. While learning theory has been influential in the development of highly effective treatments for anxiety disorders, such as exposure-based interventions, similar success is not seen in the case of addiction. Cue exposure treatments for AUD have produced mixed results.[42] The lack of strong empirical support for exposure-based treatments for addiction is largely explained by the overgeneralizability of the conditioned response. To that end, it is simply not feasible to devise exposure exercises that effectively target all such triggers. Nevertheless, functional analysis of behavior is commonly used to effectively identify patients' most salient drug use triggers. Likewise, behavioral techniques for coping with triggers, such as avoiding, taking time-outs, and learning refusal skills, represent important components of cognitive behavioral therapy for addiction. What is often lacking from this effective intervention, is the conceptualization of triggers as learned processes that are biologically based and may evoke the unwanted, yet learned, behavioral response of alcohol or drug use leading to relapse.

Learning theory is particularly useful for understanding the neural underpinnings of incentive salience in addiction. It contends that adaptive responses to various types of functional alteration are displayed not only at the level of single neurons but also at the synapses between neurons, hence, the term synaptic plasticity. This phenomenon has been most studied in the form of long-term potentiation, which is a process of long-lasting facilitation of neurotransmission across neurons when the synapses between them are used repeatedly under certain conditions. These processes are critical to all learning, both adaptive and maladaptive. Recognizing that much like the rewarding properties of alcohol and drugs that operate within the same neurocircuits responsible for normal reward functions in the brain, Pavlovian (or associative) learning during the development of addiction operates through the same signaling pathways subserving all other nonpathological forms of learning.[43] In clinical terms, patients have learned to form associations between triggers and alcohol or drug consumption through the same biological mechanisms that allow us to associate our favorite restaurant with food. The difference is that alcohol and drugs of abuse result in unnatural levels of dopamine that may be driving learning to an even greater degree than what would be expected by an unexpected reward (reward prediction errors [RPEs] are discussed in further detail below).

A number of neuroimaging studies have shown that the presentation of alcohol or drug cues, compared to control cues, reliably produces increases in blood flow in brain areas associated with reward (nucleus accumbens, ventral tegmental area [VTA], and insula)[44] and affect regulation (amygdala).[45] Studies have summarized the neural circuitry reliably involved in cue-reactivity.[46] It has been shown that alcohol cues elicit robust activation of limbic and prefrontal regions, including the ventral striatum, anterior cingulate cortex, and ventromedial prefrontal cortex in individuals with AUD. Compared to controls, individuals with AUD show greater activation of parietal and temporal regions, including the posterior cingulate, precuneus, and superior temporal gyrus. Cue-elicited activation of the ventral striatum was most frequently correlated with behavioral measures and reduced by treatment.[46,47] Importantly, while brain activation is correlated with the subjective experience of craving, captured via self-reports, the correlation is far from perfect and in some cases, not present at all.[48] This suggests that while craving is under conscious awareness, some of it may be subcortical in nature and perhaps inaccessible to patients. That is consistent with patient reports of being on "auto pilot" and having little awareness of their craving levels during a lapse. The use of functional magnetic resonance imaging (fMRI)-based cue reactivity holds promise for screening AUD/SUD treatments and provides a proof-of-mechanism for treatments thought to reduce the incentive salience of alcohol and drugs.[49] Ibudilast, a neuroimmune modulator, shows promise as a novel pharmacotherapy for AUD/SUD. A 2-week human laboratory study of ibudilast to reduce heavy drinking found that ibudilast, relative to placebo, reduced the odds of heavy drinking across time by 45% and also attenuated alcohol cue-elicited activation in the ventral striatum compared to placebo. Individuals who had attenuated ventral striatal activation and took ibudilast had the fewest number of drinks per drinking day in the week following the scan.[50] Importantly, medications development for addiction has focused on the attenuation of craving and a recent meta-analysis has found that drug cues and craving indicators play significant roles in drug use and relapse outcomes in clinical settings.[51]

## Reward/relief motivated behavior

One approach to classifying heterogeneity in SUD is based on the underlying motivation for substance use, namely, primarily using substances for their rewarding effects or primarily using substances for relief. The processes underlying these phenotypes reflect the three-stage addiction cycle.[52] Reward use is hypothesized to be more salient in the binge-intoxication stage, the initial stage of substance use wherein use is motivated by a substance's positive reinforcing effects.

Reward substance use is linked to dopaminergic and opioidergic signaling in the ventral striatum. Relief substance use is more salient in the withdrawal, negative affect stage, wherein individuals use substances to relieve negative emotionality. Relief substance use has been linked to dysregulation in norepinephrine and stress-related signaling in the extended amygdala, as well as dysfunction in glutamate and GABA signaling more broadly. A neuroimaging study from our laboratory aimed to examine the neural correlates of reward/relief behaviors.[53] Our results indicated that relief/habit drinkers (relief and habit groups were combined in this study due to overlap in clinical characteristics between these groups) showed greater dorsal striatal activation to visual alcohol cues than reward drinkers. However, cue-elicited ventral striatal activation did not differ significantly between groups. Further work identifying the neural correlates of reward and relief/habit drinking presents a path toward the refinement of these neuroscience-informed phenotypes with the ultimate goal of informing personalized treatments for AUD.

Our laboratory and others have developed validated scales to identify reward and relief drinkers. Mann et al.[37] and Roos et al.[54] used complex data reduction approaches to identify reward and relief drinkers using items from the Inventory of Drinking Situations (IDS). More recently, Votaw et al.[55] developed a 10-item version of the IDS, the Reward and Relief IDS, to identify four latent profiles: Low Reward/High Relief, Low Reward/Low Relief, High Reward/High Relief, and High Reward/Low Relief. Our laboratory recently developed the Reward, Relief, and Habit Drinking Scale (RRHDS), a short four-item self-report measure that can be used to identify drinking phenotypes in the clinic.[56]

There is some evidence that the reward/relief drinking phenotypes predict medication responses to naltrexone, an opioid antagonist, and acamprosate, a modulator of glutamatergic signaling. Several studies have found that primary reward drinkers respond better to naltrexone[37,38,57] compared to placebo, with a notable exception.[54] Roos et al.[54] showed that primary relief drinkers respond better to acamprosate than placebo, while Mann et al.[37] found no medication response to acamprosate in relief drinkers. Using the brief IDS to classify phenotypes, primary high-reward drinks responded better to naltrexone and acamprosate versus placebo.[55] To our knowledge, reward and relief phenotypes have not been applied to medication studies for other SUDs.

While identifying reward/relief phenotypes shows promise for precision medicine, prospective studies are needed to confirm the reward phenotypes and naltrexone response directly. These studies are feasible given that the RRHDS and the brief IDS can be used to stratify randomization into treatment groups by reward/relief phenotypes. While there is some evidence for the increased efficacy of naltrexone in reward users, suvorexant is a dual orexin antagonist that has garnered interest for addictive disorders.[58] Dense orexin projections from the lateral hypothalamus to the VTA provide neurobiological support that orexins may influence responses to rewarding stimuli, including alcohol.[59] Two ongoing clinical trials will assess suvorexant's potential as a treatment for AUD (NCT04229095) and comorbid AUD + insomnia (NCT03897062). Ghrelin receptor antagonists may also serve as a biological target for addictive disorders. In human laboratory studies, intravenous ghrelin administration has been shown to increase the urge to drink, increase alcohol self-administration, and modulate brain activity in regions involved in reward processing.[60,61] In regard to medications that may be beneficial for relief users, a 7-day human laboratory crossover trial of the neuroimmune modulator, ibudilast, showed that ibudilast improved mood during exposure to alcohol and stress cues and reduced the mood-altering and stimulant effects of alcohol among participants with more severe depressive symptoms.[62] Our laboratory has an ongoing phase 2 randomized controlled trial (RCT) testing ibudilast for AUD (NCT03594435).

## Reward prediction error

Learning through reinforcement is driven by a process through which an individual's behavior changes based on a pairing of decisions/actions and consequences. This type of learning is thought to be driven by the discrepancy between the predicted, or expected, outcome and the received outcome, termed RPE.[63,64] RPEs drive behavior toward the obtainment of reward. If an individual receives the predicted reward, they keep the prediction unchanged and engage in the same behavior to obtain the reward; however, if the individual does not receive the predicted reward, they must update their prediction, resulting in a changed behavior to obtain the reward. There are two types of RPEs: positive and negative. Positive prediction errors occur when a reward is better than predicted, whereas negative prediction errors occur when the reward is worse than predicted. Substantial evidence indicates that RPEs are neurally driven by dopaminergic signaling in the midbrain. Activity in midbrain dopamine neurons increases during a positive RPE and decreases during a negative RPE.

As SUDs are often defined by compulsive behavior, or use which occurs despite the experience of negative consequences, there has been substantial translational interest in RPE signaling in the field.[65] Moreover, addictive drugs have direct and indirect effects on dopaminergic neuronal activity, resulting in an increase in dopamine firing which may drive learning through an RPE-like mechanism.[65] A number of substances, including cocaine, amphetamine, alcohol, opioids, and nicotine, all acutely increase dopamine transients in the nucleus accumbens,[66,67] providing support for the theory that RPE-like signaling may drive addictive behavior.

Human studies of RPE have largely relied on fMRI paradigms to probe the neural underpinnings of the RPE signal and to examine the relationship between this signal and decision-making behavior. While fMRI task specifics differ between studies, the methodology relies on dynamic changes in response-outcome contingencies to induce RPEs.[68] In a simple version of an RPE task, a participant would be presented with two abstract stimuli. The participant would select between the stimuli and receive positive or negative feedback or a reward, typically a monetary reward. Initially, one stimulus is associated with more frequent experiences of a positive reward (i.e., 80% chance of reward), while the other stimulus is paired with a low likelihood of receiving the positive reward (i.e., 20% chance). After a participant

learns these reward contingencies, the contingencies would be unpredictably switched, such that the more frequently rewarded stimulus would become the less frequently rewarded stimulus, resulting in a negative RPE. A number of studies have used these paradigms to investigate RPE in individuals with an SUD. One of the first studies in this area found that chronic nicotine smokers produced a robust RPE signal in the caudate, a region of the dorsal striatum; however, their behavior was not guided by this neural signal.[69] Moreover, when smokers were given nicotine, neural signals guided behavior; thereby indicating that during acute abstinence and withdrawal, neural prediction error signals are intact yet are unable to influence decision-making, mirroring use despite consequences. Similarly, men with AUD displayed an intact striatal RPE signal in the brain; however, these participants did show altered functional connectivity between the striatum and the dorsolateral prefrontal cortex which was associated with impairments in learning.[70] Another study in individuals with AUD used real alcohol rewards, rather than monetary rewards, and found that individuals with AUD had increased positive RPE activity in parietal and occipital regions relative to social drinkers.[71] A recent study in adolescents found that AUD symptomology was negatively correlated with optimal decision-making and striatal and prefrontal cortex RPE signal,[72] indicating that these errors in decision-making are present in those who may not yet have a severe presentation of the disorder. Findings have been mixed in individuals with a cocaine use disorder. Initial studies found a reduced RPE signal in the striatum and orbitofrontal cortex,[73] and a reduced electroencephalogram (EEG) signal during RPEs.[74] However, another study found that during cocaine deprivation, individuals with a cocaine use disorder had heightened neural RPE activity in the striatum, which mediated the relationship between chronicity of substance use and the desire to use cocaine.[74] Overall, a recent meta-analysis of 28 studies of RPE studies in substance-using populations found that while substance users robustly activate the striatum and insula during RPEs, they show blunted activation in the putamen, inferior frontal gyrus, and right insula, relative to controls.[75] Differences in findings between and across substance-using populations may also be attributed to the satiation state of the population, (i.e., acute abstinence, active use, and acute use[76]) and by belief or expectation state.[77,78]

While understanding the neural mechanisms underlying RPEs and decision-making in SUDs holds promise for furthering treatment and medication development, few studies have directly used RPE tasks or models to evaluate pharmacological or psychosocial treatments. Instead, results from these studies provide targets for the development of learning-based addiction treatments. One medication that may hold promise to treat decision-making deficits is modafinil. Modafinil is a mild stimulant and cognitive enhancer which acts on several neurotransmitter systems, including dopamine. Modafinil has been investigated as a pharmacotherapy for AUD, methamphetamine use disorder, and cocaine use disorder. In individuals with AUD, acute treatment with modafinil improved impulsive decision-making, enhanced frontostriatal connectivity,[79] and increased response inhibition for a subset of compromised participants.[80] In individuals with a methamphetamine use disorder, acute treatment with modafinil improved performance in a learning task and increased activation in frontal brain regions.[81] However, the benefits of chronic treatment with modafinil for individuals with AUD and methamphetamine use disorder remain unknown. In individuals with cocaine use disorder, acute or short-term treatment of modafinil reduced activation in the VTA, attenuated self-reported craving,[82] and improved working memory.[83] However, an RCT of modafinil to treat cocaine use disorder failed to find improvements in clinical or neuropsychological outcomes and did not improve abstinence rates,[84] suggesting that modafinil may not be a useful treatment for cocaine use disorder. Frontal-striatal and insula-striatal pathways have been consistently found to be disrupted in RPE studies;[70,73,75] therefore, targeting these circuits may hold promise to treat SUDs.

## Goal-directed choice, habit, and compulsivity

Over the last few decades, three prominent theories of addiction have characterized addiction as either excessive goal-directed choice behaviors, maladaptive habitual behaviors, or compulsive behaviors. Goal-directed behavior is characterized by actions selected based on their resulting consequences.[85] In the context of addiction, the frequency of goal-directed drug taking is determined by the expectation of reward value combined with knowledge regarding the voluntary behaviors necessary to obtain the drug.[86,87] While the initial stages of addiction are characterized by goal-directed drug-taking behaviors, habit theory suggests that the reinforcing effects of the substance strengthen the association between drug-related stimuli and the drug-seeking response.[88] That is, habit theory suggests that addiction is driven by the strength of the established stimulus–response relationship insofar as the magnitude of the reinforcer is steady. When the reinforcer is re-experienced having a lower or higher value, the stimulus–response relationship and the frequency of drug-taking behavior adjust accordingly. As such, habitual drug-taking behavior is flexible and amenable to change, specifically when the value of the reinforcer shifts. In contrast to habit theory, compulsion theory posits that the stimulus–response relationship controlling substance use cannot be modified by re-experiencing the reinforcer at a different value.[89,90] In other words, because compulsive drug-taking behavior is driven primarily by the established stimulus–response relationship, drug-taking behavior is not impacted by a loss of value in the reinforcer.

Goal-directed decision-making requires the knowledge of the association between actions and their consequences, and is "model-based," as they rely on a model of the world to determine the best course of action.[91] Conversely, habitual decision-making is more rigid and uses past reward associations to make current decisions; referred to as "model-free" as it does not require an explicit model of the world for decision-making.[92] Several tasks have been developed to investigate goal-directed and habitual decision-making. One of the most prominent and translationally valid tasks is the two-step decision-making task developed by Daw and colleagues.[93] In this sequential two-stage task, participants choose between one of two choices in each stage, leading to a rewarded or nonrewarded outcome of varying

probabilities. In the first stage, choices are associated with a transition to a likely (70%) or unlikely (30%) state, where participants make their final selection and receive feedback. In this task, model-free, or habitual, decision-making would result in the repetition of a previously rewarded action, even if the transition was unlikely; whereas model-based, or goal-directed, decision-making considers the likelihood of the transition for future decision-making. This task has been used to investigate goal-directed and habitual decision-making in rodent models of addiction,[94] individuals with AUD,[95,96] and at-risk individuals.[97,98]

The outcome-devaluation task is also commonly used in laboratory animals and humans to differentiate between goal-directed versus habit behaviors.[99] In the typical research setting, participants learn that two responses result in two different rewarding outcomes. Of these outcomes is then devalued by allowing participants to consume the outcome to satiety or giving the participant instructions that one of the outcomes is no longer available. After the devaluation, the participants are given a choice between the two responses under conditions when the outcomes are not available (i.e., extinction). If responding to the devalued outcome decreases, it suggests that the response is goal-directed and dependent on the current value of the outcome. If responding to the devalued outcome remains steady, this suggests that the response is habitual and driven by the stimulus–response relationship. While the task has been used to provide support for habit theory among individuals who use alcohol,[100] cocaine,[101] and nicotine,[102] several studies among clinical samples failed to support habit theory (see Hogarth[99] for a thorough critique of the habit literature).

To study compulsive behavior, researchers examine actual or hypothetical drug-taking behaviors after punishment or increasing costs. For example, demand tasks examine hypothetical consumption behaviors at increasing costs.[87] The intensity of demand (maximum consumption at a low price) is considered to be a relatively pure index of drug value unaffected by costs. The breakpoint represents the price at which drug consumption drops to zero and is thought to be more sensitive to the impact of price costs on the decision to consume. Compulsion theory would be supported if substance use severity was more strongly associated with breakpoint than intensity, suggesting that cost insensitivity is more important than drug value.[103] However, meta-analyses and systematic reviews of this literature have found that indices of dependence severity correlate more consistently with measures of intensity than breakpoint, suggesting that dependence is more likely to be driven by greater drug value than cost discounting.[104–106] Deficits in reversal learning have been interpreted as evidence for greater cost discounting in addiction. In reversal learning tasks, participants learn that one response has a higher payoff than an alternative choice, before these response—reward contingencies are reversed. Individuals with SUD show deficits in reversal learning despite the comparable acquisition of the initial contingencies.[107–109] One interpretation is that drug users are less sensitive to the punishment of the incorrect choice, driving the persistence of this choice. Taken together, these studies provide minimal support for the compulsion theory.

These behavioral tasks have not been widely utilized in the context of treatments for addiction. However, behavioral and pharmacotherapies that reduce drug value may be effective at reducing goal-directed and habitual behaviors (assuming that re-experiencing the drug at a lower value will adjust the stimulus–response relationship). Effective pharmacotherapies that have been shown to decrease drug/alcohol demand include naltrexone[110] and varenicline.[111] Once proof of concept studies have established the validity of these paradigms as predictors of medication response, the tasks can be leveraged in medication development for addiction.

## CONCLUSION

SUDs and AUDs are heterogenous conditions, as currently defined in DSM-5 diagnostic criteria. Because the clinical presentation of alcohol and SUDs can vary widely from person to person, effective treatment options are limited, and a one size fits all treatment approach is not feasible. While there have been numerous attempts to classify meaningful subtypes in addiction, they have not translated into clinical practice. Thus, there is a critical need for an efficient and clinically useful way to better understand the heterogeneity in alcohol and SUDs and to leverage that knowledge toward targeted interventions. Recent efforts focused on deep behavioral phenotyping using NIAAA's Addictions Neuroclinical Assessment and/or the National Institute of Drug Abuse's Phenotypic Battery represent recent examples of neuroscience-informed constructs aimed at elucidating the clinical neuroscience of the disorders. The neuroscience-based constructs identified using deep phenotyping batteries complement core constructs in NIMH's RDoC and NIAAA's Alcohol Addiction RDoC and have the potential to extend these research frameworks to the clinic. Once these domains have been independently validated across laboratories, their predictive utility in the human laboratory and clinical trials should be put to the test. These efforts would embed neuroscience-informed phenotyping in medication development and allow for biomarker development as well. There is a critical need to apply the phenomenology to improve treatments and to develop treatment-responsive biomarkers using the underlying biology of core phenotypes (e.g., neural cue-reactivity, magnetic resonance spectroscopy assays of neurometabolites, peripheral inflammation, stress, and pharmacogenetics). This is an important direction in the implementation of neuroscience-informed phenotyping. Moreover, there have been recent calls to amplify the neuroscience-informed assessment at the level of clinical outcomes.[112]

This review addressed several neuroscience-related constructs that are proposed to underly addiction; however, this is not intended to be an exhaustive list, and it is reasonable to assume some phenotypic overlap between constructs. Our research group has tested several of the neuroscience constructs using experimental psychopathology approaches.[113] There are several opportunities for forward and reverse translation of neuroscience constructs that are worth mentioning. In the allostatic model of addiction, recruiting more severe samples of patients with AUD may reconcile the differences between the preclinical and human literature. Additionally, reverse translating reward and relief drinking phenotypes in animal models will improve our understanding of the neurodysfunction in addiction. The latter point

would be especially useful given the prominent clinical research that reward and relief phenotypes differ in their responses to pharmacological treatments. Additional examples of translational opportunities are provided elsewhere.[114] The central recommendations from the literature reviewed herein include the following: (1) implementing neuroscience-informed phenotyping as predictors and outcomes of addiction treatment in contemporary (as compared to post-hoc) samples; (2) testing the clinical utility of the neuroscientific constructs; (3) leveraging clinical neuroscience constructs toward biomarker development; (4) refining biomarkers and behavioral predictors with potential for large-scale dissemination and implementation in clinical care settings; and (5) maintaining a healthy level of skepticism regarding the "added benefit" of neuroscience-informed phenotypes and biomarkers beyond what is currently accessible to clinicians.

In summary, neuroscience constructs hold promise for advancing precision medicine for addiction. The field is experiencing a foray into their potential for informing treatment response, with most of the research focused on the reward/relief drinking phenotypes. The deep behavioral phenotyping approaches serve as a starting point to understand heterogeneity in addiction. Once the batteries are fine-tuned and validated, it is purported that they can be leveraged to develop novel interventions that target specific neuroscience-informed mechanisms that underlie addiction. Nevertheless, the proof is in the pudding. As such, we contend that a thoughtful translational neuroscience approach should consistently aim for larger effect sizes for addiction treatments, which would result in higher uptake of evidence-based practices and overall population-level improvements in healthcare outcomes for addiction.

## AUTHOR CONTRIBUTIONS

All authors significantly contributed to this manuscript. L.A.R. conceptualized and defined the objectives of the review; all authors contributed to the first draft of the manuscript and to the revision; and all authors read and edited the final manuscript and provided scientific content for it.

## ACKNOWLEDGMENTS

The authors were supported by funds from the National Institute on Alcohol Abuse and Alcoholism (K24AA025704 to L.A.R.; F32AA029288 to S.J.N.; and K01AA029712 to E.N.G.).

## COMPETING INTERESTS

The authors declare no competing interests.

## PEER REVIEW

The peer review history for this article is available at: https://publons.com/publon/10.1111/nyas.14929.

## REFERENCES

1. Project MATCH Research Group. (1997). Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, *58*, 7–29.
2. Nieto, S. J., Grodin, E. N., & Ray, L. A. (2020). On the path toward personalized medicine: Implications of pharmacogenetic studies of alcohol use disorder medications. *Expert Review of Precision Medicine and Drug Development*, *5*, 43–54.
3. Grodin, E. N., & Ray, L. A. (2019). The use of functional magnetic resonance imaging to test pharmacotherapies for alcohol use disorder: A systematic review. *Alcoholism: Clinical and Experimental Research*, *43*, 2038–2056.
4. Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*, 748–751.
5. Litten, R. Z., Ryan, M. L., Falk, D. E., Reilly, M., Fertig, J. B., & Koob, G. F. (2015). Heterogeneity of alcohol use disorder: Understanding mechanisms to advance personalized treatment. *Alcoholism: Clinical and Experimental Research*, *39*, 579–584.
6. Al-Khalil, K., Vakamudi, K., Witkiewitz, K., & Claus, E. D. (2021). Neural correlates of alcohol use disorder severity among nontreatment-seeking heavy drinkers: An examination of the incentive salience and negative emotionality domains of the alcohol and addiction research domain criteria. *Alcoholism: Clinical and Experimental Research*, *45*, 1200–1214.
7. Kwako, L. E., Momenan, R., Litten, R. Z., Koob, G. F., & Goldman, D. (2016). Addictions Neuroclinical Assessment: A neuroscience-based framework for addictive disorders. *Biological Psychiatry*, *80*, 179–189.
8. Kwako, L. E., Momenan, R., Grodin, E. N., Litten, R. Z., Koob, G. F., & Goldman, D. (2017). Addictions Neuroclinical Assessment: A reverse translational approach. *Neuropharmacology*, *122*, 254–264.
9. Kwako, L. E., Schwandt, M. L., Ramchandani, V. A., Diazgranados, N., Koob, G. F., Volkow, N. D., Blanco, C., & Goldman, D. (2019). Neurofunctional domains derived from deep behavioral phenotyping in alcohol use disorder. *American Journal of Psychiatry*, *176*, 744–753.
10. Votaw, V. R., Pearson, M. R., Stein, E., & Witkiewitz, K. (2020). The Addictions Neuroclinical Assessment negative emotionality domain among treatment-seekers with alcohol use disorder: Construct validity and measurement invariance. *Alcoholism: Clinical and Experimental Research*, *44*, 679–688.
11. Votaw, V. R., Stein, E. R., & Witkiewitz, K. (2021). A longitudinal mediation model of negative emotionality, coping motives and drinking intensity among individuals receiving community treatment for alcohol use disorder. *Alcohol Alcohol*, *56*, 573–580.
12. Stein, E. R., Votaw, V. R., Swan, J. E., & Witkiewitz, K. (2021). Validity and measurement invariance of the Addictions Neuroclinical Assessment incentive salience domain among treatment-seekers with alcohol use disorder. *Journal of Substance Abuse Treatment*, *122*, 108227.
13. Nieto, S. J., Grodin, E. N., Green, R., & Ray, L. A. (2021). Evaluation of the Addictions Neuroclinical Assessment (ANA) framework through deep phenotyping of problem drinkers. *Drug and Alcohol Dependence*, *221*, 108603.
14. Nieto, S. J., & Ray, L. A. (2022). Applying the Addictions Neuroclinical Assessment to derive neurofunctional domains in individuals who use methamphetamine. *Behavioural Brain Research*, *427*, 113876.
15. Koob, G. F., & Moal, M. L. (1997). Drug abuse: Hedonic homeostatic dysregulation. *Science*, *278*, 52–58.
16. Koob, G. F. (2003). Alcoholism: Allostasis and beyond. *Alcoholism: Clinical and Experimental Research*, *27*, 232–243.
17. Koob, G. F., & Schulkin, J. (2019). Addiction and stress: An allostatic view. *Neuroscience and Biobehavioral Reviews*, *106*, 245–262.
18. Solomon, R. L., & Corbit, J. D. (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychological Review*, *81*, 119–145.
19. Ahmed, S. H., & Koob, G. F. (2005). Transition to drug addiction: A negative reinforcement model based on an allostatic decrease in reward function. *Psychopharmacology*, *180*, 473–490.

20. Koob, G., & Kreek, M. J. (2007). Stress, dysregulation of drug reward pathways, and the transition to drug dependence. *American Journal of Psychiatry*, *164*, 1149–1159.
21. Koob, G. F., & Le Moal, M. (2008). Addiction and the brain antireward system. *Annual Review of Psychology*, *59*, 29–53.
22. Weiss, F., Parsons, L. H., Schulteis, G., Hyytiä, P., Lorang, M. T., Bloom, F. E., & Koob, G. F. (1996). Ethanol self-administration restores withdrawal-associated deficiencies in accumbal dopamine and 5-hydroxytryptamine release in dependent rats. *Journal of Neuroscience*, *16*, 3474–3485.
23. Bujarski, S., & Ray, L. A. (2014). Subjective response to alcohol and associated craving in heavy drinkers vs. alcohol dependents: An examination of Koob's allostatic model in humans. *Drug and Alcohol Dependence*, *140*, 161–167.
24. Bujarski, S., Hutchison, K. E., Prause, N., & Ray, L. A. (2015). Functional significance of subjective response to alcohol across levels of alcohol exposure. *Addiction Biology*, *22*, 235–245. https://pubmed.ncbi.nlm.nih.gov/26256114/%23
25. Bujarski, S., Jentsch, J. D, Roche, D. J. O., Ramchandani, V. A., Miotto, K., & Ray, L. A. (2018). Differences in the subjective and motivational properties of alcohol across alcohol use severity: Application of a novel translational human laboratory paradigm. *Neuropsychopharmacology*, *43*, 1891–1899.
26. King, A. C., Hasin, D., O'Connor, S. J., & Mcnamara, P. J., Cao, D. (2016). A prospective 5-year re-examination of alcohol response in heavy drinkers progressing in alcohol use disorder. *Biological Psychiatry*, *79*, 489–498.
27. Ray, L. A., Bujarski, S., Yardley, M. M., Roche, D. J. O., & Hartwell, E. E. (2017). Differences between treatment-seeking and non-treatment-seeking participants in medication studies for alcoholism: Do they matter? *American Journal of Drug and Alcohol Abuse*, *43*, 703–710.
28. Rohn, M. C. H., Lee, M. R., Kleuter, S. B., Schwandt, M. L., Falk, D. E., & Leggio, L. (2017). Differences between treatment-seeking and nontreatment-seeking alcohol-dependent research participants: An exploratory analysis. *Alcoholism: Clinical and Experimental Research*, *41*, 414–420.
29. Ray, L. A., Grodin, E. N., Leggio, L., Bechtholt, A. J., Becker, H., Feldstein Ewing, S. W., Jentsch, J. D., King, A. C., Mason, B. J., O'Malley, S., MacKillop, J., Heilig, M., & Koob, G. F. (2020). The future of translational research on alcohol use disorder. *Addiction Biology*, *26*, e12903. https://pubmed.ncbi.nlm.nih.gov/32286721/
30. Ray, L., Chin, P., & Miotto, K. (2010). Naltrexone for the treatment of alcoholism: Clinical findings, mechanisms of action, and pharmacogenetics. *CNS & Neurological Disorders Drug Targets*, *9*, 13–22.
31. Roche, D. J., & Ray, L. A. (2015). Subjective response as a consideration in the pharmacogenetics of alcoholism treatment. *Pharmacogenomics*, *16*, 721–736.
32. Ray, L. A., & Hutchison, K. E. (2007). Effects of naltrexone on alcohol sensitivity and genetic moderators of medication response: A double-blind placebo-controlled study. *Archives of General Psychiatry*, *64*, 1069–1077.
33. Ray, L. A., & Hutchison, K. E. (2004). A polymorphism of the mu-opioid receptor gene (OPRM1) and sensitivity to the effects of alcohol in humans. *Alcoholism: Clinical and Experimental Research*, *28*, 1789–1795.
34. Oslin, D. W., Berrettini, W. H., & O'Brien, C. P. (2006). Targeting treatments for alcohol dependence: The pharmacogenetics of naltrexone. *Addiction Biology*, *11*, 397–403.
35. Oslin, D. W., Leong, S. H., Lynch, K. G., Berrettini, W., O'Brien, C. P., Gordon, A. J., & Rukstalis, M. (2015). Naltrexone vs placebo for the treatment of alcohol dependence: A randomized clinical trial. *JAMA Psychiatry*, *72*, 430–437.
36. Anton, R. F., Oroszi, G., O'Malley, S., Couper, D., Swift, R., Pettinati, H., & Goldman, D. (2008). An evaluation of mu-opioid receptor (OPRM1) as a predictor of naltrexone response in the treatment of alcohol dependence: Results from the Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (COMBINE) study. *Archives of General Psychiatry*, *65*, 135–144.
37. Mann, K., Roos, C. R., Hoffmann, S., Nakovics, H., Leménager, T., Heinz, A., & Witkiewitz, K. (2018). Precision medicine in alcohol dependence: A controlled trial testing pharmacotherapy response among reward and relief drinking phenotypes. *Neuropsychopharmacology*, *43*, 891–899.
38. Witkiewitz, K., Roos, C. R., Mann, K., & Kranzler, H. R. (2019). Advancing precision medicine for alcohol use disorder: Replication and extension of reward drinking as a predictor of naltrexone response. *Alcoholism: Clinical and Experimental Research*, *43*, 2395–2405.
39. Haass-Koffler, C. L., Swift, R. M., & Leggio, L. (2018). Noradrenergic targets for the treatment of alcohol use disorder. *Psychopharmacology*, *235*, 1625–1634.
40. Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'Liking', 'wanting', and learning. *Current Opinion in Pharmacology*, *9*, 65–73.
41. Robinson, T. E., & Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction*, *96*, 103–114.
42. Conklin, C. A., & Tiffany, S. T. (2002). Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*, *97*, 155–167.
43. Robinson, T. E., & Berridge, K. C. (2008). Review. The incentive sensitization theory of addiction: Some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 3137–3146.
44. Filbey, F. M., Claus, E., Audette, A. R., Niculescu, M., Banich, M. T., Tanabe, J., Du, Y. P., & Hutchison, K. E. (2008). Exposure to the taste of alcohol elicits activation of the mesocorticolimbic neurocircuitry. *Neuropsychopharmacology*, *33*, 1391–1401.
45. Childress, A. R., Mozley, P. D, Mcelgin, W., Fitzgerald, J., Reivich, M., & O'Brien, C. P. (1999). Limbic activation during cue-induced cocaine craving. *American Journal of Psychiatry*, *156*, 11–18.
46. Schacht, J. P., & Anton, R. F., Myrick, H. (2013). Functional neuroimaging studies of alcohol cue reactivity: A quantitative meta-analysis and systematic review. *Addiction Biology*, *18*, 121–133.
47. Courtney, K. E., Schacht, J. P., Hutchison, K., Roche, D. J. O., & Ray, L. A. (2016). Neural substrates of cue reactivity: Association with treatment outcomes and relapse. *Addiction Biology*, *21*, 3–22.
48. Filbey, F. M., Schacht, J. P., Myers, U. S., Chavez, R. S., & Hutchison, K. E. (2009). Marijuana craving in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 13016–13021.
49. Grodin, E. N., & Ray, L. A. (2019). The use of functional magnetic resonance imaging (fMRI) to test pharmacotherapies for alcohol use disorders: A systematic review. *Alcoholism: Clinical and Experimental Research*, *43*, 2038–2056.
50. Grodin, E. N., Bujarski, S., Towns, B., Burnette, E., Nieto, S., Lim, A., Lin, J., Miotto, K., Gillis, A., Irwin, M. R., Evans, C., & Ray, L. A. (2021). Ibudilast, a neuroimmune modulator, reduces heavy drinking and alcohol cue-elicited neural activation: A randomized trial. *Translational Psychiatry*, *11*, 355.
51. Vafaie, N., & Kober, H. (2022). Association of drug cues and craving with drug use and relapse: A systematic review and meta-analysis. *JAMA Psychiatry*, *79*, 641–650.
52. Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology*, *35*, 217–238. https://pubmed.ncbi.nlm.nih.gov/19710631/
53. Burnette, E. M., Grodin, E. N., Schacht, J. P., & Ray, L. A. (2021). Clinical and neural correlates of reward and relief drinking. *Alcoholism: Clinical and Experimental Research*, *45*, 194–203.
54. Roos, C. R., Mann, K., & Witkiewitz, K. (2017). Reward and relief dimensions of temptation to drink: Construct validity and role in predicting differential benefit from acamprosate and naltrexone. *Addiction Biology*, *22*, 1528–1539.

55. Votaw, V. R., Mann, K., Kranzler, H. R., Roos, C. R., Nakovics, H., & Witkiewitz, K. (2022). Examining a brief measure and observed cutoff scores to identify reward and relief drinking profiles: Psychometric properties and pharmacotherapy response. *Drug and Alcohol Dependence*, *232*, 109257.
56. Grodin, E. N., Bujarski, S., Venegas, A., Baskerville, W.-A., Nieto, S. J., Jentsch, J. D., & Ray, L. A. (2019). Reward, relief and habit drinking: Initial validation of a Brief Assessment Tool. *Alcohol Alcohol*, *54*, 574–583.
57. Roos, C. R., Bold, K. W., Witkiewitz, K., Leeman, R. F., Demartini, K. S., Fucito, L. M., Corbin, W. R., Mann, K., Kranzler, H. R., & O'malley, S. S. (2021). Reward drinking and naltrexone treatment response among young adult heavy drinkers. *Addiction*, *116*, 2360–2371.
58. Campbell, E. J., Marchant, N. J., & Lawrence, A. J. (2020). A sleeping giant: Suvorexant for the treatment of alcohol use disorder? *Brain Research*, *1731*, 145902.
59. Korotkova, T. M., Sergeeva, O. A., Eriksson, K. S., Haas, H. L., & Brown, R. E. (2003). Excitation of ventral tegmental area dopaminergic and nondopaminergic neurons by orexins/hypocretins. *Journal of Neuroscience*, *23*, 7–11.
60. Farokhnia, M., Grodin, E. N., Lee, M. R., Oot, E. N., Blackburn, A. N., Stangl, B. L., Schwandt, M. L., Farinelli, L. A., Momenan, R., Ramchandani, V. A., & Leggio, L. (2018). Exogenous ghrelin administration increases alcohol self-administration and modulates brain functional activity in heavy-drinking alcohol-dependent individuals. *Molecular Psychiatry*, *23*, 2029–2038.
61. Leggio, L., Zywiak, W. H., Fricchione, S. R., Edwards, S. M., De La Monte, S. M., Swift, R. M., & Kenna, G. A. (2014). Intravenous ghrelin administration increases alcohol craving in alcohol-dependent heavy drinkers: A preliminary investigation. *Biological Psychiatry*, *76*, 734–741.
62. Ray, L. A., Bujarski, S., Shoptaw, S., Roche, D. J., Heinzerling, K., & Miotto, K. (2017). Development of the neuroimmune modulator ibudilast for the treatment of alcoholism: A randomized, placebo-controlled, human laboratory trial. *Neuropsychopharmacology*, *42*, 1776–1788.
63. Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1–27.
64. Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, *18*, 23–32.
65. Keiflin, R., & Janak, P. H. (2015). Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron*, *88*, 247–263.
66. Covey, D. P., Roitman, M. F., & Garris, P. A. (2014). Illicit dopamine transients: Reconciling actions of abused drugs. *Trends in Neurosciences*, *37*, 200–210.
67. Wanat, M., Willuhn, I., Clark, J., & Phillips, P. (2009). Phasic dopamine release in appetitive behaviors and drug addiction. *Current Drug Abuse Reviews*, *2*, 195–213.
68. Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *Journal of Cognitive Neuroscience*, *21*, 1332–1345.
69. Chiu, P. H., Lohrenz, T. M., & Montague, P. R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nature Neuroscience*, *11*, 514–520.
70. Park, S. Q., Kahnt, T., Beck, A., Cohen, M. X., Dolan, R. J., Wrase, J., & Heinz, A. (2010). Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. *Journal of Neuroscience*, *30*, 7749–7753.
71. Cservenka, A., Courtney, K. E., Ghahremani, D. G., Hutchison, K. E., & Ray, L. A. (2017). Development, initial testing and challenges of an ecologically valid reward prediction error FMRI task for alcoholism. *Alcohol Alcohol*, *52*, 617–624.
72. Aloi, J., Crum, K. I., Blair, K. S., Zhang, R., Bashford-Largo, J., Bajaj, S., Schwartz, A., Carollo, E., Hwang, S., Leiker, E., Filbey, F. M., Averbeck, B. B., Dobbertin, M., & Blair, R. J. R. (2021). Individual associations of adolescent alcohol use disorder versus cannabis use disorder symptoms in neural prediction error signaling and the response to novelty. *Developmental Cognitive Neuroscience*, *48*, 100944.
73. Tanabe, J., Reynolds, J., Krmpotich, T., Claus, E., Thompson, L. L., Du, Y. P., & Banich, M. T. (2013). Reduced neural tracking of prediction error in substance-dependent individuals. *American Journal of Psychiatry*, *170*, 1356–1363.
74. Parvaz, M. A., Konova, A. B., Proudfit, G. H., Dunning, J. P., Malaker, P., Moeller, S. J., Maloney, T., Alia-Klein, N., & Goldstein, R. Z. (2015). Impaired neural response to negative prediction errors in cocaine addiction. *Journal of Neuroscience*, *35*, 1872–1879.
75. Tolomeo, S., Yaple, Z. A., & Yu, R. (2021). Neural representation of prediction error signals in substance users. *Addiction Biology*, *26*, e12976.
76. Baker, T. E., Zeighami, Y., Dagher, A., & Holroyd, C. B. (2020). Smoking decisions: Altered reinforcement learning signals induced by nicotine state. *Nicotine & Tobacco Research*, *22*, 164–171.
77. Wang, G.-J., Wiers, C. E., Shumay, E., Tomasi, D., Yuan, K., Wong, C. T., Logan, J., Fowler, J. S., & Volkow, N. D. (2019). Expectation effects on brain dopamine responses to methylphenidate in cocaine use disorder. *Translational Psychiatry*, *9*, 93.
78. Gu, X., Lohrenz, T., Salas, R., Baldwin, P. R., Soltani, A., Kirk, U., Cinciripini, P. M., & Montague, P. R (2015). Belief about nicotine selectively modulates value and reward prediction error signals in smokers. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 2539–2544.
79. Schmaal, L., Goudriaan, A. E., Joos, L., Dom, G., Pattij, T., van den Brink, W., & Veltman, D. J. (2014). Neural substrates of impulsive decision making modulated by modafinil in alcohol-dependent patients. *Psychological Medicine*, *44*, 2787–2798.
80. Schmaal, L., Joos, L., Koeleman, M., Veltman, D. J., Van Den Brink, W., & Goudriaan, A. E. (2013). Effects of modafinil on neural correlates of response inhibition in alcohol-dependent patients. *Biological Psychiatry*, *73*, 211–218.
81. Ghahremani, D. G., Tabibnia, G., Monterosso, J., Hellemann, G., Poldrack, R. A., & London, E. D. (2011). Effect of modafinil on learning and task-related brain activity in methamphetamine-dependent and healthy individuals. *Neuropsychopharmacology*, *36*, 950–959.
82. Goudriaan, A. E., Veltman, D. J., Van Den Brink, W., Dom, G., & Schmaal, L. (2013). Neurophysiological effects of modafinil on cue-exposure in cocaine dependence: A randomized placebo-controlled cross-over study using pharmacological fMRI. *Addictive Behaviors*, *38*, 1509–1517.
83. Kalechstein, A. D., Mahoney, J. J., Yoon, J. H., Bennett, R., & De La Garza, R. (2013). Modafinil, but not escitalopram, improves working memory and sustained attention in long-term, high-dose cocaine users. *Neuropharmacology*, *64*, 472–478.
84. Karila, L., Leroy, C., Dubol, M., Trichard, C., Mabondo, A., Marill, C., Dubois, A., Bordas, N. G., Martinot, J.-L., Reynaud, M., & Artiges, E. (2016). Dopamine transporter correlates and occupancy by modafinil in cocaine-dependent patients: A controlled study with high-resolution PET and [(11)C]-PE2I. *Neuropsychopharmacology*, *41*, 2294–2302.
85. Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464–476.
86. Müller, C. P., & Schumann, G. (2011). Drugs as instruments: A new framework for non-addictive psychoactive drug use. *Behavioral and Brain Sciences*, *34*, 293–310.
87. Mackillop, J. (2016). The behavioral economics and neuroeconomics of alcohol use disorders. *Alcoholism: Clinical and Experimental Research*, *40*, 672–685.

88. Dickinson, A., Wood, N., & Smith, J. W. (2002). Alcohol seeking by rats: Action or habit? *Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, *55*, 331–348.
89. Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*, 1481–1489.
90. Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, *67*, 23–50.
91. Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*, 407–419.
92. Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704–1711.
93. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204–1215.
94. Groman, S. M., Massi, B., Mathias, S. R., Lee, D., & Taylor, J. R. (2019). Model-free and model-based influences in addiction-related behaviors. *Biological Psychiatry*, *85*, 936–945.
95. Sebold, M., Deserno, L., Nebe, S., Schad, D. J., Garbusow, M., Hägele, C., Keller, J., Jünger, E., Kathmann, N., Smolka, M., Rapp, M. A., Schlagenhauf, F., Heinz, A., & Huys, Q. J. M. (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*, *70*, 122–131.
96. Sebold, M., Nebe, S., Garbusow, M., Guggenmos, M., Schad, D. J., Beck, A., Kuitunen-Paul, S., Sommer, C., Frank, R., Neu, P., Zimmermann, U. S., Rapp, M. A., Smolka, M. N., Huys, Q. J. M., Schlagenhauf, F., & Heinz, A. (2017). When habits are dangerous: Alcohol expectancies and habitual decision making predict relapse in alcohol dependence. *Biological Psychiatry*, *82*, 847–856.
97. Chen, H., Mojtahedzadeh, N., Belanger, M. J., Nebe, S., Kuitunen-Paul, S., Sebold, M., Garbusow, M., Huys, Q. J. M., Heinz, A., Rapp, M. A., & Smolka, M. N. (2021). Model-based and model-free control predicts alcohol consumption developmental trajectory in young adults: A 3-year prospective study. *Biological Psychiatry*, *89*, 980–989.
98. Obst, E., Schad, D. J., Huys, Q. J., Sebold, M., Nebe, S., Sommer, C., Smolka, M. N., & Zimmermann, U. S. (2018). Drunk decisions: Alcohol shifts choice from habitual towards goal-directed control in adolescent intermediate-risk drinkers. *Journal of Psychopharmacology*, *32*, 855–866.
99. Hogarth, L. (2020). Addiction is driven by excessive goal-directed drug choice under negative affect: Translational critique of habit and compulsion theory. *Neuropsychopharmacology*, *45*, 720–735.
100. Sjoerds, Z., De Wit, S., Van Den Brink, W., Robbins, T. W., Beekman, A. T. F., Penninx, B. W. J. H., & Veltman, D. J. (2013). Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Translational Psychiatry*, *3*, e337.
101. Ersche, K. D., Gillan, C. M., Jones, P. S, Williams, G. B., Ward, L. H. E., Luijten, M., De Wit, S., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2016). Carrots and sticks fail to change behavior in cocaine addiction. *Science*, *352*, 1468–1471.
102. Luijten, M. (2019). *Goal-directed and habitual control in smokers*. Radboud University.
103. Murphy, J. G., & Mackillop, J. (2006). Relative reinforcing efficacy of alcohol among college student drinkers. *Experimental and Clinical Psychopharmacology*, *14*, 219–227.
104. Zvorsky, I., Nighbor, T. D., Kurti, A. N., Desarno, M., Naudé, G., Reed, D. D., & Higgins, S. T. (2019). Sensitivity of hypothetical purchase task indices when studying substance use: A systematic literature review. *Preventive Medicine*, *128*, 105789.
105. González-Roz, A., Jackson, J., Murphy, C., Rohsenow, D. J., & Mackillop, J. (2019). Behavioral economic tobacco demand in relation to cigarette consumption and nicotine dependence: A meta-analysis of cross-sectional relationships. *Addiction*, *114*, 1926–1940.
106. González-Roz, A., Secades-Villa, R., Weidberg, S., García-Pérez, Á., & Reed, D. D. (2020). Latent structure of the cigarette purchase task among treatment-seeking smokers with depression and its predictive validity on smoking abstinence. *Nicotine & Tobacco Research*, *22*, 74–80.
107. Ersche, K. D., Roiser, J. P., Robbins, T. W., & Sahakian, B. J. (2008). Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. *Psychopharmacology*, *197*, 421–431.
108. Reiter, A. M. F., Deserno, L., Kallert, T., Heinze, H.-J., Heinz, A., & Schlagenhauf, F. (2016). Behavioral and neural signatures of reduced updating of alternative options in alcohol-dependent patients during flexible decision-making. *Journal of Neuroscience*, *36*, 10935.
109. Vanes, L. D., Holst, R. J., Jansen, J. M., Brink, W., Oosterlaan, J., & Goudriaan, A. E. (2014). Contingency learning in alcohol dependence and pathological gambling: Learning and unlearning reward contingencies. *Alcoholism: Clinical and Experimental Research*, *38*, 1602–1610.
110. Bujarski, S., Mackillop, J., & Ray, L. A. (2012). Understanding naltrexone mechanism of action and pharmacogenetics in Asian Americans via behavioral economics: A preliminary study. *Experimental and Clinical Psychopharmacology*, *20*, 181–190.
111. Green, R., & Ray, L. A. (2018). Effects of varenicline on subjective craving and relative reinforcing value of cigarettes. *Drug and Alcohol Dependence*, *188*, 53–59.
112. Goldstein, R. Z. (2022). Neuropsychoimaging measures as alternatives to drug use outcomes in clinical trials for addiction. *JAMA Psychiatry*, *79*, 843.
113. Ray, L. A., & Grodin, E. N. (2021). Clinical neuroscience of addiction: What clinical psychologists need to know and why. *Annual Review of Clinical Psychology*, *17*, 465–493.
114. Nieto, S. J., Grodin, E. N., Aguirre, C. G., Izquierdo, A., & Ray, L. A. (2021). Translational opportunities in animal and human models to study alcohol use disorder. *Translational Psychiatry*, *11*, 496.


**How to cite this article:** Ray, L. A., Nieto, S. J., & Grodin, E. N. (2023). Translational models of addiction phenotypes to advance addiction pharmacotherapy. *Ann NY Acad Sci.*, *1519*, 118–128. https://doi.org/10.1111/nyas.14929