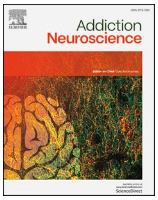
Contents lists available at ScienceDirect

Addiction Neuroscience

journal homepage: www.elsevier.com/locate/addicn



The utility of a latent-cause framework for understanding addiction phenomena

Sashank Pisupati [a,d], Angela J. Langdon [b], Anna B. Konova [c], Yael Niv [*,d]

[a] *Limbic Limited, London UK*
[b] *National Institute of Mental Health & National Institute on Drug Abuse, National Institutes of Health, Bethesda MD, USA*
[c] *Department of Psychiatry, University Behavioral Health Care & Brain Health Institute, Rutgers University, New Brunswick NJ, USA*
[d] *Princeton Neuroscience Institute & Department of Psychology, Princeton University, Princeton NJ, USA*

ARTICLE INFO

Keywords:
Latent-cause inference
Addiction
Craving
Relapse

ABSTRACT

Computational models of addiction often rely on a model-free reinforcement learning (RL) formulation, owing to the close associations between model-free RL, habitual behavior and the dopaminergic system. However, such formulations typically do not capture key recurrent features of addiction phenomena such as craving and relapse. Moreover, they cannot account for goal-directed aspects of addiction that necessitate contrasting, model-based formulations. Here we synthesize a growing body of evidence and propose that a latent-cause framework can help unify our understanding of several recurrent phenomena in addiction, by viewing them as the inferred return of previous, persistent “latent causes”. We demonstrate that applying this framework to Pavlovian and instrumental settings can help account for defining features of craving and relapse such as outcome-specificity, generalization, and cyclical dynamics. Finally, we argue that this framework can bridge model-free and model-based formulations, and account for individual variability in phenomenology by accommodating the memories, beliefs, and goals of those living with addiction, motivating a centering of the individual, subjective experience of addiction and recovery.

Introduction: addiction and model-free reinforcement learning

Addiction is often defined as a chronic, relapsing condition[1] characterized by compulsive behaviors that continue despite adverse consequences. As a result, many prominent computational models conceptualize addiction as an aberrant[2] regime of the brain’s learning and decision-making systems [3,4]. In particular, early computational accounts of addiction focused on the effects of addictive substances on trial-and-error learning, or so-called model-free reinforcement learning (RL; Box 1), and the resulting hijacking of habitual behavior. There is a growing appreciation of alternative accounts that focus on deliberative planning/decision-making and model-based RL (see [4–6] for in-depth discussions of this literature), however we focus here on model-free RL as the most well-characterized account.

One popular model-free RL account suggests that the unnaturally reinforcing properties of addictive substances give rise to features of addiction such as an inflated valuation of cues signaling drug rewards above and beyond those signaling non-drug rewards, and compulsive (i.e., hard to resist) habitual behaviors resulting from such overvaluation [7–9]. This view offers principled ways to model the effects of addiction observed in value-based decisions or economic choices [5,10–12]. It also explains connections between addiction and the neuromodulator dopamine, which is a major molecular target of many addictive substances, and is thought to be a critical element of model-free RL algorithms such as temporal-difference (TD) learning in the brain – namely, a reward prediction error signal that is crucial for learning.

For instance, addictive drugs have been hypothesized to give rise to a persistent, non-compensable version of the prediction error in TD

* Corresponding author.
E-mail address: yael@princeton.edu (Y. Niv).

[1] Note: We have deliberately eschewed the use of the word “disorder” here, in an effort to use non-pathologizing language around addiction, while acknowledging its disabling nature and recognizing that different individuals relate differently to their addiction

[2] Such models often implicitly assume that this “aberration” lies within the brain, but many authors have questioned this view, arguing for agnosticism and/or careful consideration of social determinants e.g. [1]. In our reformulation (similar to [2]), we attempt to take a biopsychosocial view of a system operating as usual, but under unusual circumstances.

https://doi.org/10.1016/j.addicn.2024.100143
Received 7 January 2023; Received in revised form 12 January 2024; Accepted 12 January 2024
Available online 15 January 2024

learning algorithms that learn state values, leading to the overvaluation of drug-predictive cues [13]. In variants of TD learning that learn state-action values, extended drug exposure has been proposed to alter the basal reward level relative to which prediction errors are calculated, to account for the development of compulsive drug-taking actions and increased impulsivity [14]. Finally, several extensions of these models of addiction rely on actor-critic formulations, which use TD prediction errors to simultaneously update a "critic" that learns state values and an "actor" that separately learns action policies using signals from the critic (Box 1). Proposals for the influence of addictive drugs in these models range from aberrant learning that originates in the critic and consequently affects the actor [15,16] to aberrant prediction-error signalling that affects both the critic and actor simultaneously [15,17], yielding compulsive habits.

Challenges to model-free RL formulations of addiction

Despite these successes in accounting for overvaluation and compulsion, purely model-free RL formulations fail to capture many crucial features of addiction phenomena such as outcome-specific craving (e.g. [18]), the dynamics of spiralling or relapse (e.g. [2]), and alterations in processes such as planning and goal-directed choice (e.g. [19,20]). This is due to the fact that most model-free RL formulations rely on learning a scalar value for any given state or action that abstracts away information about specific forthcoming outcomes and uses cached past experiences rather than future goals for fast evaluation. Importantly, these models treat learning and unlearning as symmetric processes. Hence, although model-free RL formulations (and the habitual and/or dopaminergic systems they are associated with) are prominent in models of addiction, several authors have suggested that there is a need for additional theory development, and a complete account must involve model-based RL systems (Box 1), which rely on a predictive model of the world, hence retaining outcome sensitivity and being capable of flexible goal-directed planning [5,21,22].

In this vein, a separate line of investigation has emerged that focuses on modeling the process of Bayesian inference of variables such as outcomes or actions under a predictive world model, and alterations of this process in addiction [4,5]. For instance, reduced behavioral and neural sensitivity to Bayesian prediction errors in a stop-signal task was shown to predict relapse in individuals with methamphetamine dependence [23]. Similarly, active-inference formulations (i.e., using a predictive model of preferred outcomes to infer the most likely actions) have demonstrated reduced sensitivity to unpleasant outcomes in substance use disorders [5]. Prominent models of craving also rely on model-based planning [24] or Bayesian inference accounts [25,26].

In an attempt to reconcile the effects of addictive substances on both model-free and model-based systems, some "dual systems" accounts of addiction invoke interactions between the two systems, such as partial model-based evaluation terminating in model-free values [19], changes in the competition between the two systems [5,12,27,28], or changes in substrates common to both such as biased samples from memory during evaluation [2] or replay [19] to account for craving, relapse and other seemingly goal-directed phenomena in addiction. Here we present a unified framework that offers explanations for several phenomena that challenge model-free formulations of addiction, and potentially bridges model-free and model-based explanations by reformulating the problem of reinforcement learning into one involving the discovery and inference of "latent causes".

Box 1. (*Basic terminology*)

Reinforcement learning (RL): A computational theory for learning from feedback (rewarding or punishing outcomes **O**) about the best actions **A** to take in an environment.

State: A representation of sensory observations **S**, including internal or external cues, that acts as a substrate for learning. The goal of reinforcement learning is to specify what action to make in what state.

Action policy: The probability of taking each action **A** in a given state **S**.

Value: A typically scalar representation of the cumulative future reward expected from a given state **V(S)** or from a particular action taken in that state **Q(S,A)**.

Model-based RL: A form of RL that uses a learned internal model of the transitions between states in the world, and the outcome in each state, to simulate the consequences of actions.

Model-free RL: A form of RL that uses past experience to learn values and/or policies through trial-and-error, without estimating a model of the world.

Instrumental learning: A learning setting in which outcomes **O** are contingent on performing certain actions **A** in a given state **S.**

Pavlovian learning: A learning setting in which outcomes **O** are predicted by the presence of a certain state **S** alone, without any action contingencies.

Latent cause: A grouping of sensory observations according to a hypothesized shared hidden variable ”generating” the sensory observations. The structure and presence of latent (hidden) causes must be **inferred**.

The latent-cause framework

A growing body of evidence has demonstrated the shortcomings of traditional model-free formulations of addiction, and has attempted to remedy them in multiple, seemingly distinct ways. However, one principle that emerges from these attempts and unifies them is the "contextual" or "latent-cause framework". According to the latent-cause framework, individuals engaging in real world reinforcement learning are not simply learning unitary associations between reinforcers and their antecedent cues or actions (such as values or policies in traditional model-free RL), but are simultaneously discovering the hidden "contexts" or "latent causes" underlying both cues and reinforcing outcomes in the environment [29,30]. This means that individuals are constantly segmenting their experiences into (an unknown number of) distinct latent causes, such that experiences only create or modify associations bound to the latent cause they are assigned to. Depending on the level of abstraction at which they are inferred, latent causes may correspond to different psychological constructs. These may be external to the individual - ranging from "categories" e.g. a *desirable object* to an entire "context" such as a *stressful situation* - or internal, from discrete phenomenological "states" such as *craving* to an entire self-concept e.g. the *addicted version of me*.

Formally, such a process of latent cause discovery can be captured using Bayesian nonparametric models [29] with "infinite-capacity" priors that allow agents to flexibly create new latent causes in the face of unfamiliar or changing experiences, store old latent causes in memory, and infer their recurring presence when familiar experiences are encountered. Importantly, latent causes are unobserved, learnt in an unsupervised fashion, and beliefs about their inferred presence are heavily shaped by one's priors and past experiences such that two individuals with differing histories may arrive at different latent cause structures underlying the same set of experienced cues and reinforcing outcomes [31]. This flexibility in learnt latent structure sets these models apart from standard Bayesian inference accounts that typically assume a single, known latent structure, and allows these models to additionally capture the process of building a world model in the first place. Beyond addiction, such a framework has successfully accounted

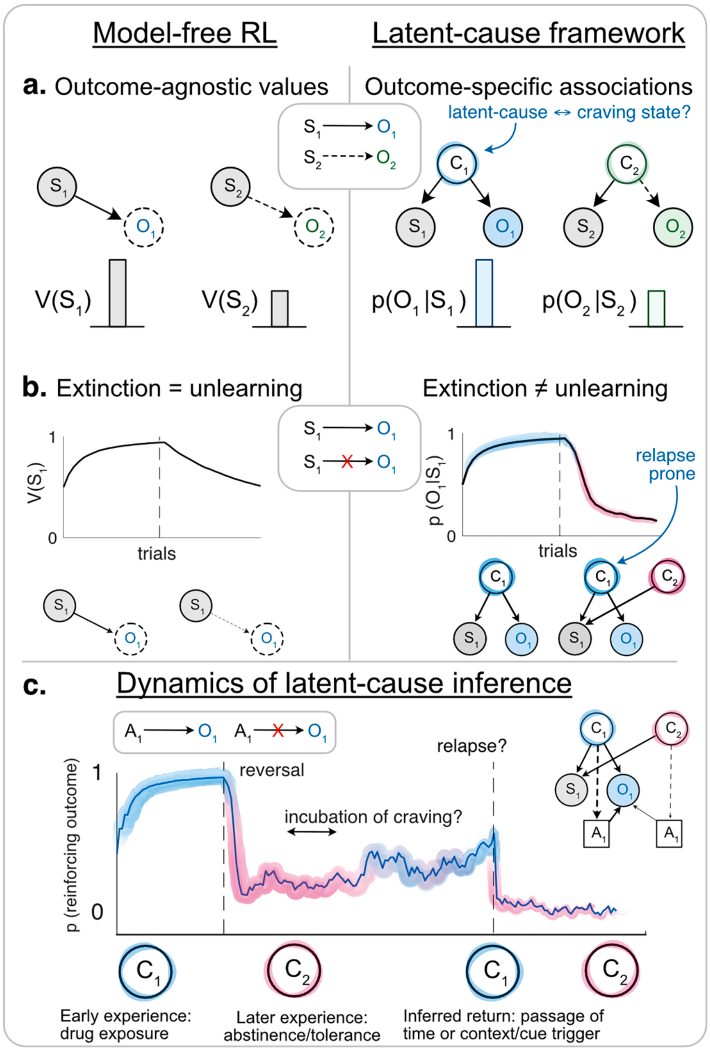

Model-free RL
Latent-cause framework
a. Outcome-agnostic values
Outcome-specific associations
latent-cause ↔ craving state?
S1 → O1
S2 → O2
C1
C2
S1
O1
S2
O2
V(S1)
V(S2)
p(O1 |S1)
p(O2 |S2)
b. Extinction = unlearning
Extinction ≠ unlearning
V(S1)
1
0
trials
p (O1|S1)
relapse prone
c.
Dynamics of latent-cause inference
A1 → O1
p (reinforcing outcome)
reversal
incubation of craving?
relapse?
Early experience: drug exposure
Later experience: abstinence/tolerance
Inferred return: passage of time or context/cue trigger

Fig. 1. The latent-cause framework and its utility in understanding craving and relapse. a-b. Differences between model-free RL (left column) and the latent-cause framework (right column). **a. Outcome specificity of associations and craving states**: Two outcome identities, O_1 and O_2, are paired with stimuli S_1 and S_2, respectively (middle). Left: Model-free RL abstracts away the outcome identities in favor of scalar values V (grey bars). Right: latent causes labelled C_1 and C_2 also represent abstractions, however, these are of both stimuli (cues) and outcomes, and retain outcome-specific expectations (colored bars) despite being “defocused” and generalizable, much like craving states. **b. Asymmetry of learning and unlearning and vulnerability to relapse**: In extinction, O_1 that was initially paired with S_1, no longer appears (middle). Left: In model-free RL, learning and unlearning are symmetric, such that the value of S_1 first increases (when paired with O_1) and then decreases in extinction (the first trial of extinction is marked by a dotted line). Right: In the latent-cause framework, the large change between acquisition and extinction can entail the formation of a new latent cause (pink) that is associated only with S_1 and not with O_1, leaving previously learnt knowledge (blue) intact, and therefore prone to relapse. **c. Dynamics of latent-cause inference, craving and relapse**: Simulation of the timecourse of latent-cause inference in reversal learning, showing potential connections to addiction phenomena. Early experience with an addictive substance leads to robust learning of drug associations and a strong policy of choosing the consumption action A_1 (square), bound to a latent cause with strong expectations of reinforcing outcomes (blue). Following reversal (first dashed line), the act of consumption no longer gives rise to the same level of reinforcement, either due to tolerance (reduction in the hedonic value of the drug outcome) or other negative consequences of drug use, leading to abstinence and a drop in outcome expectations. This sudden change leads to the creation of a new latent cause (pink) that is not as strongly bound to the action policy or its outcome, and this new cause persists for some time. As time passes, exhaustion of this persistence or exposure to drug-related cues or contextual triggers increase the probability of inferring the return of the original drug-associated cause (blue). This leads to an increase in outcome expectations, resembling incubation of craving. Eventually, outcome expectations and/or craving may grow so large as to trigger a relapse event (second dashed line), resetting the cycle.

for a wide variety of learning phenomena in humans and animals [30], and its neural underpinnings are beginning to be dissected in the prefrontal cortex and hippocampus [32–34].

The latent-cause framework deviates from standard model-free RL in a number of key ways: First, it puts cues and outcomes on the same level, and learns latent causes associated with both (Fig. 1a), hence retaining information about specific outcomes and enabling both cues and outcomes to trigger memories or expectancies of each other, and of old latent causes, any of which could act as potential substrates of outcome-specific craving [18,24,35]. Second, it allows for the segregation of associations, values or policies into multiple separate representations (one for each latent cause), enabling the learning of distinct (and even conflicting) associations for different contexts (Fig. 1b). Hence, initial drug exposure could create learning of one set of associations, while later experiences such as spiraling into addiction, adverse outcomes, development of tolerance or abstinence from the drug could create a separate set of associations rather than modify the original association. The resultant protection of the original association leaves the individual vulnerable to relapse if and when the original latent cause is inferred to recur [2,36]. Similar ideas have been proposed and tested in the domain of fear learning and extinction [37] with applications to fear-related psychopathologies [38]. Finally, early protected latent causes could be sampled from memory in the future during decision evaluation or replay, leading to goal-directed behavior such as "chasing the first high" [2] and fully deliberative, model-based control that is nevertheless insensitive to devaluation [39]. In the following two sections, we introduce and briefly examine the consequences of the latent-cause framework for Pavlovian and instrumental learning, and its potential to explain core addiction phenomena that is not well captured by pure model-free RL accounts.

Pavlovian latent causes and craving

When applied to Pavlovian learning settings, the latent-cause framework has the potential to explain aspects of craving that model-free RL cannot. Since in this framework, individuals associate latent causes with both cues and reinforcing outcomes (rather than learn scalar values for cues in model-free RL), both cues and outcomes have the capacity to trigger inference of the associated latent cause, and consequently, expectations of the other stimulus. Importantly, since outcome information is retained (rather than abstracted away), these expectations are *outcome-specific* (Fig. 1a). Latent states formed by drug-related cues and outcomes could potentially capture many properties of the subjective state of craving. The inferred, transient, presence of these latent states can be triggered by drug-associated or other contextual cues. In turn, the inferred states can give rise to outcome expectations and inflated values that are specific to both the drug outcome and the temporally restricted state of craving [18]. Moreover, since latent causes represent groupings of cues, outcomes and any surrounding context, they are expected to generalize to stimuli that resemble those cues and outcomes. Such restricted similarity-based generalization of drug-related cues and outcomes has been observed in states of craving [18,40], and is particularly evident in polydrug users whose cravings for specific drugs can be triggered by historically drug-associated moods [41]. This is also akin to the "defocusing" of outcome representations proposed by [42], which can enable the substitution of one addictive outcome with a similar one.

The dynamics of learning and inference about latent causes may help shed further light on temporal features of craving – for instance, "incubation of craving" [26,43,44], i.e. the phenomenon that craving reduces immediately after consumption but may increase steadily after a period of abstinence, with a characteristic timescale. Over the longer term, craving may continue to recur frequently (and with it the risk of relapse) or become more intermittent and even decrease, with longer adherence to abstinence. Such dynamics could potentially be captured by the distinct but linked timescales of inference and learning in the latent-cause framework (Fig. 1c). After abrupt change points, one might infer that transient latent causes have become inactive. The anticipation of their short-term return may increase gradually with the probability of the current transient latent cause terminating (e.g., after a temporary period of abstinence) or increase abruptly in the presence of a drug-related cue, outcome or antecedent context (e.g. a stressful event, [26]). However, the long-term recurrence of these latent causes may be determined by learnt priors over latent causes, which in turn may depend on their "popularity" – a latent cause that has been sampled from memory more (or less) often in the past may be expected to return more (or less) often in the future [2], acquiring a higher (or lower) learnt prior probability. Understanding individual differences in such priors and other components of learning and inference could potentially help disambiguate situations in which craving continues to recur and accelerate relapse from those in which it eventually reduces and leads to recovery [2,26]. It is worth noting here that while we are attempting to model the overall dynamics of craving in terms of an unobservable belief state similar to [26], empirical measures of craving differ significantly between human (self-report) and rodent (responding in extinction) studies, and appropriate care will be needed in order to map our predictions about latent craving states onto these vastly different observable measures.

Instrumental latent causes and relapse

When applied to instrumental learning scenarios, the latent-cause framework can account for the asymmetric dynamics of learning and unlearning of drug-related associations, and the phenomenon of relapse [2,36]. Akin to the situation-recognition module proposed in [36], the process of latent-cause inference has the potential to segregate learning during initial acquisition of a behavioral policy (Box 1) (e.g. initial drug consumption with large experienced payoffs) from learning during subsequent extinction or reversal of outcomes (e.g. abstinence, worsening payoffs due to tolerance or increasing costs), and can explain why both drug consumption and rehabilitation may be less likely to generalize outside the context they were learnt in [45]. Segregation of different experiences to separate latent causes makes old associations resistant to unlearning or counter-evidence, and prone to relapse (Fig. 1b). This can also explain behavior of "chasing the first high" despite the lack of recent experience of positive outcomes of drug consumption [2], and accounts not only for the return of old policies (i.e. early drug consumption) but also their associated contextual memories (i.e. the first high). As in the Pavlovian case, such relapses (i.e., inference of resurgence of an old latent cause) can be triggered by drug-associated cues, passively re-experiencing drug outcomes (i.e. reinstatement, [46]), or antecedent contexts (e.g. a stressful state that prompted initial consumption), and may be accompanied by outdated action-outcome expectations. Finally, latent causes associated with old policies may or may not be the same latent causes associated with outcomes, allowing for the possibility of craving and relapse to exist without each other [47], while still making one likely in the presence of the other – another potential axis of individual variability.

Latent causes as a bridge between model-free and model-based formulations

So far, we have considered the effect of latent-cause representations underlying values and policies that are still learnt through direct experience, similar to model-free associations. Going one step further, latent-causes could serve as building blocks of a predictive model of the world, acting as higher-order, partially observable state representations upon which model-based evaluation could take place. For instance, a drug-related cue could cause inference of a latent cause that effectively puts an individual into a state of craving or relapse, triggering either an old model-free habit, an outdated model-based evaluation [39] or a goal to alleviate craving and a model-based search in pursuit of this goal [24]. In this way, the latent-cause framework acts as a bridge between model-free and model-based formulations, offering a common substrate for both habitual and goal-directed aspects of addiction [19,21,47]. By bridging these model classes, the latent cause framework could potentially provide new links between results couched in terms of model-free RL and the growing body of work on the bidirectional influences of substance use on model-based planning and goal-directed decision making [4–6].

Looking ahead: centering subjective experience

Beyond its ability to account for a number of features of addiction phenomena, the latent-cause framework shifts the emphasis away from experience-derived model-free learning that "takes control" over behavior and thus seemingly removes agency from the individual, instead putting the focus on subjective aspects of that experience and how it may interact with the past to influence decision-making in context [1,48,49]. Under this view, exteroceptive and interoceptive experiences are filtered through an individual's model of latent causes in the world, which have been shaped by their past experiences, prior beliefs (both innate and learned) and goals. For instance, a person who expects – as a prior on how the world works – that latent causes are deterministic (also called a “black or white” cognitive bias, where events can be good or bad, safe or dangerous), would be more likely to infer segregated latent causes across conflicting experiences, which in turn can yield update-resistant beliefs or policies that may later resurface [50]. Such prior expectations can be due to both innate tendencies and life experiences, and may comprise a self-reinforcing cycle as inference in light of prior beliefs can update prior expectations for the future. This account therefore leaves room for – and can help explain – individual differences in phenomenology and responding to the same set of drug-related or therapeutic experiences, due to differences in individual histories and innate tendencies. Moreover, it emphasizes the importance and therapeutic potential of samples from past memories as well as alternate imagined futures, in particular predicting that strengthening alternate drug-unrelated memories and futures is likely to be more successful at supporting recovery than attempting to erase or curb drug-related ones [2].

Finally, the latent-cause framework can be applied to inference at many different levels of abstraction—all the way up to the individual’s own sense of self and agency, by treating the self-concept as a latent cause that groups together the characteristics, values and action policies of an individual to distinguish themselves from the environment. Individual differences in inferring self v.s. environmental latent causes could yield different (and even conflicting) narratives of identity in addiction, from individuals conceptualizing their self as a persistent whole that goes through addiction and possibly recovery, to split senses of self (an addicted and a non-addicted identity with conflicting values and policies), to addiction as being a cause that is entirely extraneous to identity [51,52]. The expressive power of this framework offers the potential for future work to attempt modeling complex, personalized phenomena in addiction within a single framework, all the while recognizing the insufficiency and limitations of unified narratives [1,53].

CRediT authorship contribution statement

Sashank Pisupati: Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Angela J. Langdon:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Anna B. Konova:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Yael Niv:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

We would like to thank Sama Ahmed, Nathaniel Daw, Gary Aston Jones, Diksha Gupta, Kirsten Smith, and David Epstein for helpful discussions and feedback on this work. This research was supported by NIDA grants R01DA054201 and R01DA042065, NIMH grant R01MH119511, a grant from the Templeton Foundation and the Intramural Research Program of the NIMH.

References

[1] H. Pickard, The puzzle of addiction. The Routledge Handbook of Philosophy and Science of Addiction, Routledge, 2018, pp. 9–22.

[2] A.M. Bornstein, H. Pickard, ǣchasing the first highg̱: memory sampling in drug choice, Neuropsychopharmacology 45 (6) (2020) 907–915.

[3] B. Gutkin, S.H. Ahmed, Computational Neuroscience of Drug Addiction vol. 10, Springer Science & Business Media, 2011.

[4] J.A. Mollick, H. Kober, Computational models of drug use and addiction: a review, J. Abnormal Psychol. 129 (6) (2020) 544.

[5] R. Smith, S. Taylor, E. Bilek, Computational mechanisms of addiction: recent evidence and its relevance to addiction medicine, Curr. Addict. Rep. (2021) 1–11.

[6] M.C. Gueguen, E.M. Schweitzer, A.B. Konova, Computational theory-driven studies of reinforcement learning and decision-making in addiction: what have we learned? Curr. Opin. Behav. Sci. 38 (2021) 40–48.

[7] Z. Kurth-Nelson, A.D. Redish, Modeling decision-making systems in addiction. Computational Neuroscience of Drug Addiction, Springer, 2012, pp. 163–187.

[8] M. Keramati, A. Dezfouli, P. Piray, Understanding addiction as a pathological state of multiple decision making processes: a neurocomputational perspective. Computational Neuroscience of Drug Addiction, Springer, 2012, pp. 205–233.

[9] Q.J. Huys, L. Deserno, K. Obermayer, F. Schlagenhauf, A. Heinz, Model-free temporal-difference learning and dopamine in alcohol dependence: examining concepts from theory and animals in human imaging, Biol. Psychiatry Cognit. Neurosci. Neuroimaging 1 (5) (2016) 401–410.

[10] W.Y. Ahn, J. Dai, J. Vassileva, J.R. Busemeyer, J.C. Stout, Computational modeling for addiction medicine: from cognitive models to clinical applications, Prog. Brain Res. 224 (2016) 53–65.

[11] T.V. Lim, K.D. Ersche, Theory-driven computational models of drug addiction in humans: fruitful or futile? Addict. Neurosci. 5 (2023) 100066.

[12] J. Vassileva, J.-H. Lee, E. Psederska, W.-Y. Ahn, Utility of computational approaches for precision psychiatry: applications to substance use disorders. Computational Neuroscience, Springer, 2023, pp. 211–231.

[13] A.D. Redish, Addiction as a computational process gone awry, Science 306 (5703) (2004) 1944–1947.

[14] A. Dezfouli, P. Piray, M.M. Keramati, H. Ekhtiari, C. Lucas, A. Mokri, A neurocomputational model for cocaine addiction, Neural Comput. 21 (10) (2009) 2869–2893.

[15] Y. Takahashi, G. Schoenbaum, Y. Niv, Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model, Front. Neurosci. (2008) 14.

[16] P. Piray, M.M. Keramati, A. Dezfouli, C. Lucas, A. Mokri, Individual differences in nucleus accumbens dopamine receptors predict development of addiction-like behavior: a computational approach, Neural Comput. 22 (9) (2010) 2334–2368.

[17] P. Dayan, Dopamine, reinforcement learning, and addiction, Pharmacopsychiatry 42 (S 01) (2009) S56–S65.

[18] A.B. Konova, K. Louie, P.W. Glimcher, The computational form of craving is a selective multiplication of economic value, Proc. Natl. Acad. Sci. 115 (16) (2018) 4122–4127.

[19] D.A. Simon, N.D. Daw, Dual-system learning models and drugs of abuse. Computational Neuroscience of Drug Addiction, Springer, 2012, pp. 145–161.

[20] L. Hogarth, Addiction is driven by excessive goal-directed drug choice under negative affect: translational critique of habit and compulsion theory, Neuropsychopharmacology 45 (5) (2020) 720–735.

[21] A.D. Redish, S. Jensen, A. Johnson, Addiction as vulnerabilities in the decision process, Behav. Brain Sci. 31 (4) (2008) 461–487.

[22] D.H. Epstein, Let’s agree to agree: a comment on Hogarth (2020), with a plea for not-so-competing theories of addiction, Neuropsychopharmacology 45 (5) (2020) 715–716.

[23] K.M. Harlé, J.Y. Angela, M.P. Paulus, Bayesian computational markers of relapse in methamphetamine dependence, NeuroImage Clin. 22 (2019) 101794.

[24] A.D. Redish, A. Johnson, A computational model of craving and obsession.(2007).

[25] X. Gu, F. Filbey, A Bayesian observer model of drug craving, JAMA Psychiatry 74 (4) (2017) 419–420.

[26] X. Gu, Incubation of craving: a Bayesian account, Neuropsychopharmacology 43 (12) (2018) 2337–2339.

[27] K.M. Fraser, P.H. Janak, How does drug use shift the balance between model-based and model-free control of decision making? Biol. Psychiatry 85 (11) (2019) 886–888.

[28] M. Sebold, L. Deserno, S. Nebe, D.J. Schad, M. Garbusow, C. Hägele, J. Keller, E. Jünger, N. Kathmann, M. Smolka, et al., Model-based and model-free decisions in alcohol dependence, Neuropsychobiology 70 (2) (2014) 122–131.

[29] S.J. Gershman, D.M. Blei, Y. Niv, Context, learning, and extinction, Psychol. Rev. 117 (1) (2010) 197.

[30] S.J. Gershman, K.A. Norman, Y. Niv, Discovering latent causes in reinforcement learning, Curr. Opin. Behav. Sci. 5 (2015) 43–50.

[31] S. Pisupati, A. Langdon, Y. Niv, Two factors underlying maladaptive inference of causal structure can drive resistance to extinction in anxiety, Biol. Psychiatry 89 (9) (2021) S283.

[32] J.B. Heald, D.M. Wolpert, M. Lengyel, The computational and neural bases of context-dependent learning, Annu. Rev. Neurosci. 46 (2023).

[33] E.D. Boorman, S.C. Sweigart, S.A. Park, Cognitive maps and novel inferences: a flexibility hierarchy, Curr. Opin. Behav. Sci. 38 (2021) 141–149.

[34] A.R. Vaidya, D. Badre, Abstract task representations for inference and control, Trends Cognit. Sci. (2022).

[35] S.T. Tiffany, Cognitive concepts of craving, Alcohol Res. Health 23 (3) (1999) 215.

[36] A.D. Redish, S. Jensen, A. Johnson, Z. Kurth-Nelson, Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling, Psychol. Rev. 114 (3) (2007) 784.

[37] J.E. Dunsmoor, Y. Niv, N. Daw, E.A. Phelps, Rethinking extinction, Neuron 88 (1) (2015) 47–63.

[38] A. Norbury, H. Brinkman, M. Kowalchyk, E. Monti, R.H. Pietrzak, D. Schiller, A. Feder, Latent cause inference during extinction learning in trauma-exposed individuals with and without PTSD, Psychol. Med. 52 (16) (2022) 3834–3845.

[39] N. Garrett, S. Allan, N.D. Daw, Model based control can give rise to devaluation insensitive choice, bioRxiv (2022).

[40] K. Biernacki, S. Lopez-Guzman, J.C. Messinger, N.V. Banavar, J. Rotrosen, P. W. Glimcher, A.B. Konova, A neuroeconomic signature of opioid craving: How fluctuations in craving bias drug-related and nondrug-related value, Neuropsychopharmacology 47 (8) (2022) 1440–1448.

[41] D.H. Epstein, J. Willner-Reid, M. Vahabzadeh, M. Mezghanni, J.-L. Lin, K. L. Preston, Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use, Arch. Gen. Psychiatry 66 (1) (2009) 88–94.

[42] P. Dayan, K.C. Berridge, Model-based and model-free pavlovian reward learning: revaluation, revision, and revelation, Cognit. Affect. Behav. Neurosci. 14 (2) (2014) 473–492.

[43] J.W. Grimm, B.T. Hope, R.A. Wise, Y. Shaham, Incubation of cocaine craving after withdrawal, Nature 412 (6843) (2001) 141–142.

[44] C.L. Pickens, M. Airavaara, F. Theberge, S. Fanous, B.T. Hope, Y. Shaham, Neurobiology of the incubation of drug craving, Trends Neurosci. 34 (8) (2011) 411–420.

[45] L.N. Robins, J.E. Helzer, D.H. Davis, Narcotic use in southeast asia and afterward: an interview study of 898 vietnam returnees, Arch. Gen. Psychiatry 32 (8) (1975) 955–961.

[46] M. Song, C.E. Jones, M.-H. Monfils, Y. Niv, Explaining the effectiveness of fear extinction through latent-cause inference, arXiv preprint arXiv:2205.04670(2022).

[47] A.D. Redish, Implications of the multiple-vulnerabilities theory of addiction for craving and relapse, Addiction 104 (11) (2009) 1940–1941.

[48] X. Gu, T.H. FitzGerald, K.J. Friston, Modeling subjective belief states in computational psychiatry: interoceptive inference as a candidate framework, Psychopharmacology 236 (8) (2019) 2405–2412.

[49] C.-H. Kao, G. Feng, J.K. Hur, H. Jarvis, R. Rutledge, Computational models of subjective feelings in psychiatry, 2022. psyarxiv.com/kq8vf. 10.31234/osf.io/kq8vf.

[50] S. Pisupati, Y. Niv, Why do some beliefs and action policies resist updating?. The 5th Multidisciplinary Conference on Reinforcement Learning and Decision Making, 2022, p. P495.

[51] J. McIntosh, N. McKeganey, Addicts' narratives of recovery from drug use: constructing a non-addict identity, Soc. Sci. Med. 50 (10) (2000) 1501–1510.

[52] H. Pickard, Addiction and the self, Noûs 55 (4) (2021) 737–761.

[53] R.R. Hammer, M.J. Dingel, J.E. Ostergren, K.E. Nowakowski, B.A. Koenig, The experience of addiction as told by the addicted: incorporating biological understandings into self-story, Cult. Med. Psychiatry 36 (4) (2012) 712–734.