Computational Mechanisms of the Dual-System Imbalance between Goal-Directed and Habitual Control in Substance Use Disorders

Authors: Yufeng Xia, Ruyuan Zhang, Ruyuan Zhang

Date: 2026-04-17T18:43:42+00:00

Abstract

The core characteristic of substance use disorders is the compulsive seeking and intake of addictive substances, a process associated with the imbalance between the dual systems of goal-directed and habit control. This review systematically organizes three computational models used to explain the dual-system imbalance in substance use disorders: the arbiter model, the hierarchical control model, and the successor representation model. Among these, the arbiter model currently occupies the mainstream in research. This review summarizes the core computational methods of each model and the current status of their applications in substance use disorder research. Future research should simultaneously incorporate the three computational mechanisms for comparative validation to more comprehensively reveal the causes underlying the formation of compulsive behavior in substance use disorders.

Full Text

Preamble

Dual-System Imbalance Between Goal-Directed and Habitual Control in Substance Use Disorders Yufeng Xia, Ruyuan Zhang

(Institute of Brain Health, National Center for Mental Disorders, Shanghai Mental Health Center and School of Psychology, Shanghai Jiao Tong University School of Medicine, Shanghai, 200030)

Individuals with substance use disorders (SUD) often struggle to establish effective behavioral goals, leading to decision-making that disregards consequences and results in compulsive drug-seeking and consumption behaviors. This review systematically organizes the computational methods and related empirical studies used to analyze this phenomenon. We summarize the conclusions

and limitations of these approaches and provide corresponding reference code. This synthesis facilitates a deeper understanding of the developmental processes underlying substance use disorders. Furthermore, the relevant computational principles and code provided herein are broadly applicable to other fields of cognitive neuroscience research.

Abstract

The core characteristic of substance use disorders (SUD) is the compulsive seeking and consumption of addictive substances, a process closely associated with an imbalance between the goal-directed and habit-control systems. This review systematically organizes three computational models used to explain this dual-system imbalance in SUD: the arbiter model, the hierarchical control model, and the successor representation model. Among these, the arbiter model currently dominates the research landscape. This review summarizes the core computational methods of each model and their current applications in the study of substance use disorders. Future research should incorporate all three computational mechanisms for comparative validation to more comprehensively reveal the underlying causes of compulsive behavior in substance use disorders.

Keywords

Substance use disorders, goal-directed, habit control, compulsive behavior

CLC Number: B846

** Corresponding author: Ruyuan Zhang, Male, Associate Professor at Peking University, e-mail: ruyuanzhang@gmail.com.

1 Introduction

Substance use disorder (SUD) is a psychiatric condition characterized by high prevalence, severe physical and psychological impairment, and significant social harm [?, ?, ?]. A hallmark symptom of SUD is an addictive craving for substances, manifesting as compulsive desires and behaviors directed toward the intake of addictive substances such as drugs, alcohol, and nicotine [?, ?, ?]. This compulsive behavior refers to an individual' s persistent seeking and consumption of substances despite the clear knowledge that such actions will lead to severe negative consequences. Compulsivity directly predicts the severity of substance abuse in addicted individuals and is regarded by researchers as a core attribute of the addiction concept [?, ?, ?].

Numerous studies have demonstrated that before substance intake transitions into compulsive behavior, it often evolves from controllable goal-directed behavior to habitual control behavior [?, ?, ?, ?]. Specifically, goal-directed behavior refers to actions taken by an individual based on an accurate assessment of behavioral outcomes and their values; such behaviors are typically highly flexible but require greater cognitive effort. In contrast, habitual control behavior refers to responses repeatedly triggered by environmental cues, characterized by relatively fixed and rigid patterns [?, ?, ?, ?, ?]. In the early stages of substance use disorder, individuals are still able to make relatively flexible substance-use decisions based on their needs and subjective goals. However, with repeated substance use, some individuals begin to rely on prior reinforcement experiences and environmental cues, gradually developing habitual substance use.

Although an increase in habitual control behavior does not inevitably lead to compulsive substance seeking, some habitual behaviors eventually evolve into compulsive seeking and intake when accompanied by abnormalities in neural circuits such as the prefrontal-striatal pathway [?, ?]. It can be argued that the behavioral shift from goal-directed to habitual control is often a precursor to compulsive substance use. Therefore, gaining a deep understanding of the process underlying this transition is key to understanding the development of substance use disorders.

Substantial evidence reveals that this transition process is related to an imbalance between the dual systems of goal-directed and habitual control [?, ?, ?, ?]. In healthy populations, the goal-directed system guides individuals to learn the associations between behaviors and goals, allowing them to master the structure of the environment and tasks so that goal values can be updated promptly when circumstances change [?, ?]. As learning and training progress and individuals become fully familiar with the task structure, they gradually rely on the habitual control system for learning [?, ?]. However, when negative outcomes or feedback occur, they rapidly revert to a goal-directed learning mode [?, ?, ?], maintaining a dynamic balance between the two. Conversely, extensive evidence suggests that chronic substance abuse is associated with a significant disruption of this dynamic balance, potentially leading to an increase in habitual control behaviors [?, ?, ?, ?, ?, ?]. For example, even when the association between task goals and rewards has been degraded, individuals who abuse cocaine continue to respond habitually to goals, indicating a failure to adjust behavioral strategies in response to changes in goals [?, ?, ?]. Similar results have been validated in individuals with nicotine and alcohol use disorders [?, ?]. Even in the early stages before substance dependence has fully formed, dual-system imbalance may already constitute a significant risk factor for predicting subsequent substance use [?, ?]. However, an increase in habitual control behavior does not inherently imply a dual-system imbalance, as habitual responding can also be an adaptive and efficient strategy under conditions of high training intensity or high environmental stability [?, ?]. More precisely, the critical impact of dynamic dual-system imbalance lies not simply in generating more habitual responses, but in increasing an individual' s reliance on established habitual response patterns, making it more difficult to adjust behavior in a timely manner based on goal values or environmental changes [?, ?, ?]. In such cases, individuals are more likely to exhibit persistent dependence on addictive substances during subsequent development, creating the conditions for the formation of compulsive behavior.

While numerous studies have indicated that dual-system imbalance is a key mechanism in the formation of compulsive behavior in substance use disorders, there is currently considerable debate regarding whether this imbalance stems primarily from impairment of the goal-directed system or an over-enhancement of the habitual control system. Substance use disorders are indeed accompanied by the emergence of more habitual behaviors. For a long time, many perspectives have suggested that over-enhanced habitual control leads to this phenomenon [?, ?]. Theoretically, however, an abnormality in either of the two systems could disrupt their original dynamic balance. Consequently, some researchers have proposed that this phenomenon does not directly prove a problem with the habitual control system itself. Instead, it may result from impaired functioning of the goal-directed system, where individuals find it difficult to flexibly adjust their behavior based on outcomes [?, ?]. Recent reviews also suggest that the goal-directed and habitual control systems may operate in parallel in various ways, making their relative contributions to behavior difficult to measure; thus, the association between substance use disorders and the habitual control system requires cautious evaluation [?, ?, ?, ?]. These controversies limit researchers' understanding of dual-system imbalance and the mechanisms of compulsivity in substance use disorders.

Although it remains difficult to clearly distinguish the specific source of dual-system imbalance in substance use disorders based on behavioral performance alone, computational models based on reinforcement learning theory provide a more precise means of exploration. By decomposing the underlying computational parameters of behavior, computational models help further identify the sources of dual-system imbalance. Currently, using computational models to characterize the relative contributions of goal-directed and habitual control systems has gradually become an important research path for revealing how substance use behavior shifts toward habituation and eventually evolves into compulsive behavior [?, ?]. However, research in this field still faces several challenges. First, in current studies of substance use disorders, findings exploring dual-system imbalance using mainstream computational models are not entirely consistent [?, ?]. Second, some researchers argue that current mainstream computational models and related tasks may have significant limitations in predicting individual changes in the dual systems [?, ?]. Finally, with the development of reinforcement learning theory itself, researchers have proposed new computational theories or models for assessing the relationship between the two systems [?, ?, ?]. The research value of these theories and models also warrants further consideration. Therefore, to further elucidate the process of dual-system imbalance in substance use disorders, it is necessary to systematically review various current and emerging computational mechanisms.

Based on the above background, this review aims to provide a detailed synthesis of current research on the computational mechanisms underlying the imbalance between goal-directed and habitual control systems in substance use disorders. Based on different assumptions regarding the relationship between the two systems, this review introduces three computational models

Figure 1

Figure 1: Figure 1

. Specifically, this review focuses on summarizing the findings and limitations of the mainstream "arbitrator" model in substance use disorder research, while also introducing two computational models—based on hierarchical control and successor representation mechanisms—that hold exploratory value for the future of the field. These conclusions will help clarify the computational principles of dual-system imbalance in substance use disorders and elucidate the possible mechanisms of compulsive substance seeking and intake. This review hopes that these conclusions will inspire future computational research in the field and provide a theoretical basis for developing new quantitative clinical indicators targeting the compulsive symptoms of substance use disorders.

2 Classical Computational Methods for Dual Systems

The computational processes underlying goal-directed and habitual control systems originated from two classes of methods in reinforcement learning theory: model-based (MB) and model-free (MF) approaches. Specifically, MB methods correspond to the goal-directed system, while MF methods correspond to the habitual control system. Numerous studies in both behavioral and neural sciences support a consistent mapping between the MB/MF framework and these dual systems [**?**, **?**, **?**, **?**].

The MB approach assumes that an agent possesses an internal model of the environment or task structure. This model incorporates state-transition probabilities—the likelihood of reaching a subsequent state given a specific action taken in a current state. By utilizing this internal model, the agent can predict and plan corresponding behavioral sequences, enabling it to select actions that maximize rewards based on defined goals; thus, this method corresponds to the goal-directed system. A typical computational process for the MB approach is as follows:

$$Q_{MB}(s,a) = \sum_{s'} P(s'|s,a)V(s')$$

In Model-Based (MB) methods, the agent learns by calculating the state prediction error (SPE) and updating the state transition probability matrix. The update mechanism is typically integrated into the iterative optimization framework. Specifically, the updated transition probability a_{ij} is defined as:

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

where $\xi_t(i,j)$ represents the probability of being in state i at time t and state j at time $t+1$. By continuously updating the state transition probability matrix, the model can more accurately reflect the underlying temporal dependencies.

In contrast to MB methods, Model-Free (MF) methods do not maintain an internal model of the environment. Instead, the agent directly updates the expected reward value of taking a specific action in its current state based solely on immediate reward feedback. Consequently, this approach corresponds to the habitual control system. A typical computational process for an MF method, such as SARSA, is as follows:

$$Q_{MF}(s_t,a_t) \leftarrow Q_{MF}(s_t,a_t) + \alpha[r_{t+1} + \gamma Q_{MF}(s_{t+1},a_{t+1}) - Q_{MF}(s_t,a_t)]$$

where α is the learning rate and γ is the discount factor.

3.1 Computational Methods of the Arbitrator Model

To fully characterize the imbalance between goal-directed and habitual control systems, it is first necessary to establish the nature of the relationship between these two systems. Extensive neurological research indicates that goal-directed and habitual control systems coexist within the human brain and maintain a competitive relationship characterized by mutual constraint [?, ?, ?, ?]. Based on this hypothesis, several researchers have proposed "arbitrator" models grounded in MB and MF methods [?, ?, ?]. This framework has gradually become the dominant model in the field [?, ?, ?]. The arbitrator model posits that an imbalance between the dual systems occurs when one system becomes either too weak or too dominant during competition.

The arbitrator model assumes that an agent utilizes both MB and MF methods for learning simultaneously, but an "arbitration" process determines the proportion to which each method influences the agent' s final decision. The core computational process of this model is as follows:

$$Q_{net}(s,a) = w \cdot Q_{MB}(s,a) + (1-w) \cdot Q_{MF}(s,a)$$

Where $0 \leq w \leq 1$ represents the arbitration weight. A higher w value indicates that the agent' s choices rely more heavily on the MB method (goal-directed system), while a lower value indicates greater reliance on the MF method (habitual system). The choice of action is typically calculated via the SoftMax function:

$$P(a|s) = \frac{\exp(\beta Q_{net}(s,a))}{\sum_{a'} \exp(\beta Q_{net}(s,a'))}$$

where β is the exploration parameter.

3.2 Application of Arbitrator Models in SUD Research

Studies of substance use disorders (SUD) based on the arbitrator model have yielded conflicting results. Research has primarily focused on alcohol use disorder (AUD). Early behavioral studies found that individuals with AUD relied significantly less on the goal-directed system compared to healthy controls [?, ?]. Subsequent research demonstrated that the MB weight in individuals with AUD was lower than that of healthy control groups and remained static across varying conditions, suggesting a failure to maintain a dynamic balance [?, ?, ?]. However, other studies have failed to find significant differences in the dual-system balance between healthy controls and individuals with AUD or high-risk populations [?, ?, ?].

Evidence regarding methamphetamine use disorder (MUD) remains equally controversial. One study found that both MB and MF weights decreased following methamphetamine self-administration, which does not support the dual-system imbalance hypothesis [?, ?]. Another study found that MB weights in individuals with MUD did not differ from those of healthy controls [?, ?]. Conversely, some research has indicated that individuals with MUD, binge eating disorder, and obsessive-compulsive disorder (OCD) all exhibit lower MB weights compared to controls, suggesting a shared mechanism across compulsion-related disorders [?, ?].

3.3 Limitations of Arbitrator Models

First, the application of arbitrator models in SUD research remains limited, primarily concentrated on alcohol use disorder. Second, arbitrator models have failed to resolve the core controversy regarding whether dual-system imbalance stems from goal-directed impairment or habitual enhancement. Third, the heavy reliance on the two-stage task may limit the consistency of conclusions, as task comprehension and structural settings can significantly influence MB/MF weighting [?, ?, ?].

4 Hierarchical Control Models in SUD

The arbitrator model assumes a competitive relationship, but this may oversimplify the interaction. Some researchers have proposed a hierarchical control model, arguing that the relationship between the dual systems should be collaborative [?, ?]. In this model, habit is defined as a combination of action sequences and a process of automatic execution, rather than a behavior controlled solely by a habitual system.

The hierarchical control model assumes that habitual behavior is regulated by the goal-directed system. Habitual behavior is not a single action, but an action sequence formed through learning. As training progresses, the agent treats this sequence as a single unit, executed automatically without relying on feedback from each step. This framework suggests that many behaviors previously categorized as habit-controlled may, in fact, be automated action sequences regulated by a higher-level goal-directed system.

5 Successor Representation Models in SUD

Beyond hierarchical control, researchers have proposed the successor representation (SR) model. The SR model constructs a compromise computational process situated between goal-directed and habitual control to explain the balance between flexibility and automation [?, ?, ?]. The SR model assumes that agents store the expected future occupancy of subsequent states when reaching a particular state:

$$M(s, s') = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \mathbb{I}(s_t = s') | s_0 = s\right]$$

This schema functions similarly to the transition matrix in MB methods but with reduced computational cost. The value of state s can be calculated as:

$$V(s) = \sum_{s'} M(s, s') R(s')$$

The SR model provides a new perspective on the dual-system imbalance. Erroneous or rigid successor representations can trigger habitual or compulsive behaviors, suggesting that some manifestations in SUD may not arise from abnormalities in the goal-directed or habit systems themselves, but from the formation of incorrect internal representations of the task structure.

6 Summary and Outlook

This review organizes three computational mechanisms regarding dual-system imbalance: the arbiter model, the hierarchical control model, and the successor representation model. While the arbiter model is the most widely applied, its findings in SUD remain inconsistent. Hierarchical control and SR models offer novel perspectives that challenge the traditional competitive framework. Future research should incorporate these diverse computational accounts to more comprehensively reveal the underlying causes of compulsive behavior in substance use disorders and provide a theoretical basis for developing new quantitative clinical indicators.

Figures

Source: ChinaXiv —Machine translation. Verify with original.

Figure 2

Figure 2: Figure 2